Essays in banking and international finance

Schäfer, Larissa

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Essays in Banking and International Finance

Essays in Banking and International Finance

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op maandag 2 november 2015 om 14.15 uur door

Larissa Schäfer

geboren op 29 juli 1985 te Wolgograd, Rusland.

Promotor: prof. dr. V.P. Ioannidou

Copromotor: dr. L.T.M. Baele

Overige leden van de Promotiecommissie:

dr. F. Braggion

prof. dr. F.C.J.M. de Jong

dr. A. Manconi

dr. S. Steffen

Acknowledgements

This thesis is the result of the most exciting, challenging and enlightening years of my life. I would not have been able to write this thesis, had it not been for the support from many people, whom I would like to thank now.

First of all, I would like express my sincere gratitude to my supervisors, Vasso Ioannidou and Lieven Baele, for their constant support, patience, motivation, and guidance. I could not have imagined a better mix of supervisors. Both young, motivated, enthusiastic and engaged researchers, open and willing to teach and guide me. When I first met Vasso she told me that doing research changes your way of thinking and lets you approach life differently which was exactly what I was looking for. When working with Vasso, I enjoyed our stimulating and insightful discussions and our efficient work routine. She always found the right balance between guiding me and giving me the freedom to work on my own. She taught me everything it takes to be an excellent researcher in banking.

When I met Lieven, I immediately got along with him well and was interested in working on Central and Eastern European equity markets, not least because of my Russian heritage. I enjoyed working with Lieven on such a completely different topic than banking that challenged me to think in new directions. At the same time, we could quickly get off the research topic and talk about life in general, which I also enjoyed very much. During my last year Lieven did not only give me practical advice but also moral support through the ups and downs of the job market.

My coauthors, Thorsten Beck and Geert Bekaert, also greatly contributed to my journey to become a researcher. I was fortunate to work on projects with such outstanding and experienced researchers. Working together with Thorsten was fantastic since I learned a great deal not only

about research but also on how to make it interesting for a broader audience. He and Vasso then gave me the opportunity to use propriety data from Armenia and conduct my own research for which I am thankful to both of them. I would also like to thank Geert for his elaborate, thoughtful, witty comments and attention to detail on our joint paper as well as for sponsoring me to visit Columbia University.

My deep gratitude goes to the members of the my dissertation committee, Fabio Braggion, Frank de Jong, Alberto Manconi and Sascha Steffen for taking the time to read my dissertation, for giving me valuable and insightful comments and for honoring me with their presence. I would also like to thank the current and former faculty members, Steven Ongena and Hans Degryse, as well as the secretaries of the Department of Finance at Tilburg University for their continuous support.

I would also like to thank my host at the EBRD, Ralph de Haas, for giving me insights into the BEPS II Survey, connecting me with experts on Armenia and giving me the opportunity to present my work to practitioners. During my time at Columbia University, I was lucky to meet Andrew Hertzberg who gave me invaluable comments on my job market paper and introduced me to his colleagues.

My friends and fellow PhDs created the perfect work-life balance for me. I would like to thank Anton, Bernardus, Cisil, and Paola for sticking together from the first hard year until now and for all the fun time we had together; Elisabeth and Masha, my paranymphs, for their support, encouragement, lots of laughs and good stories; my office mates Radomir and Hao for creating a nice working environment; Geraldine and Michela for being the most fun visiting researchers; Olga and Ina for keeping in touch and supporting me through all stages of my life. I would like to thank Carsten for his support during the last year, for making me laugh and forget everything, for showing me how crazy our research world for outsiders may seem.

Lastly, I would like to thank my parents for their continued love, support, encouragement and dedication. Liebe Mama, lieber Papa, wenn ihr damals nicht nach Deutschland gezogen wäret, wäre ich nie so weit gekommen. Ich danke euch herzlichst für diesen mutigen Schritt und all die Möglichkeiten, die ihr mir damit eröffnet habt. Durch eure Liebe, Unterstützung, Ermutigung und Hingabe habe ich mich stark genug gefühlt jede Aufgabe zu meistern.

Introduction

This Ph.D. dissertation consists of three chapters, two in banking and one in international finance. The first two chapters study how foreign ownership structures and bank business models affect bank-firm lending relationships. The last chapter focuses on Central and Eastern European equity market development and integration.

The first chapter examines whether foreign banks can overcome their informational disadvantages with contract design and credit scoring models. Foreign bank entry has been witnessed across the globe and especially in emerging markets. Understanding whether and how foreign banks may be able to overcome distance constraints and penetrate new and less developed markets is key for understanding the role they can play in such markets and designing supportive policies. To answer this question we explore differences in foreign and domestic banks' credit contract terms and pricing models. Using a sample of firms that borrow from both domestic and foreign banks in the same month, we show that foreign banks are more likely to demand collateral and grant shorter maturity loans than domestic banks. Foreign banks also base their pricing on internal credit ratings and collateral pledges, while domestic banks price according to the length, depth and breadth of their relationship with a firm. These findings confirm that foreign banks can overcome distance constraints using contract design and credit scoring models. Such mechanisms however may carry important limitations. The stronger reliance on collateral and short maturity loans may reduce the options for firm investment. It may also have important repercussions from a capital allocation perspective, as it implies a shift towards firms with short-term financing needs and with pledgeable assets both on the intensive and extensive margin, limiting the potentially beneficial role foreign banks can play in developing countries.

The second chapter studies whether relationship lending helps firms that are in financial distress. Understanding bank behavior when firms are hit by idiosyncratic liquidity shocks is important for practitioners and academics alike. Firms and their sources of funding when faced with liquidity shocks are vital for the economy as a whole. Moreover, insights into banks' reactions to idiosyncratic shocks to firms will help design policies to foster financial stability. Combining a new and direct measure of relationship lending with unique credit registry data, I examine the effect of relationship lending on ex-post loan performance. My findings demonstrate that the same firm is more likely to become temporary delinquent on a relationship-based relative to a transaction-based loan. Consistent with theory, relationship banks tolerate temporary bad results, yet extract rents in the long run. When firms are in distress, relationship banks adjust contract terms and offer drawdowns on credit lines and overdrafts but do not rollover loans more often. Moreover, relationship banks are more likely to continue to lend to firms after past non-performance. Overall, the paper uncovers a new channel of how relationship lending serves as a liquidity insurance for firms in distress. In the end, liquidity insurance should advance longer investment horizons of firms and thus higher employment and growth. From a policy perspective, the better understanding of bank lending mechanisms and sources of firm funding will help designing policies to foster financial stability.

The last chapter provides a comprehensive and detailed analysis of Central and Eastern European equity markets from the mid-1990s until now and evaluates the value of investing in these markets for global investors. Despite their disappointing performance over the last years, emerging markets continue to attract a lot of interest from investors. They are believed to offer higher expected returns accompanied by diversification benefits. Using firm-level data and custom-made indices and indicators, we show that (i) there is considerable heterogeneity in the degree, dynamics, and determinants of market development across the different markets, (ii) that especially the smaller markets still offer diversification benefits to global investors, and (iii) that there are substantial premiums associated to investing in small, value, illiquid and low volatility stocks.

Contents

Chapter 1

Foreigners vs. Native: Bank Lending Technologies and Loan Pricing

Coauthors: Thorsten Beck, Vasso Ioannidou

1.1 Introduction

A burgeoning literature on financial intermediation studies the impact of "distance" between borrowers and lenders in credit markets. Distance — be it geographical, cultural, institutional, or organizational — can aggravate information asymmetries and change incentives in ways that may lead to worse credit outcomes. The large increase in foreign bank entry across the globe and especially in emerging markets is one such example. A view espoused by many scholars is that "distance constraints" prevent foreign banks from effectively penetrating such markets by influencing who they can lend to and how. This paper studies how foreign banks overcome distance constraints by comparing the lending practices of domestic and foreign banks. It is novel as it focuses on a group of firms that borrow simultaneously from both domestic and foreign banks, which allows us to hold constant differences in their clienteles. Understanding whether and how foreign banks may be able to overcome distance constraints and penetrate new and less developed markets is key for understanding the role they can play in such markets and designing supportive policies.

While several studies document that foreign banks tend to "cherry pick" the largest and

most transparent firms, shying away from smaller, more opaque firms, particularly in developing countries and emerging markets, the evidence is far from conclusive.[1] Distance constraints are argued to create extra informational and enforcement costs for foreign banks, putting foreign banks at a disadvantage particularly when lending to information and monitoring-intensive firms (Mian (2006)). Distance constraints could arise because of the greater geographical distance between the bank's headquarters (principal) and the local branches (agent), differences in corporate culture, legal environment, and institutional framework between the home and the host county as well as organizational differences. Due to their typically larger size and legal structure (i.e., a parent abroad), foreign banks have more hierarchical organizational structures (i.e., more layers in the decision making) that create diseconomies in lending to smaller, more opaque firms (Stein (2002)). The maintained assumption is that lending to such firms relies heavily on "soft" information that is difficult to quantify and transmit to the higher levels of the organization, where the credit allocation decision lies. Others argue that foreign banks can overcome such constraints using credit scoring models and contract design (Berger and Udell (2006)). Credit scoring models allow banks to "harden" soft information and transmit it to the higher levels of the organization enabling its use in the credit allocation process.[2] Contractual features such as collateral and maturity can also be used to mitigate asymmetric information and enforcement problems such as adverse selection, moral hazard, and bank losses in default (see, among others, Bester (1985); Flannery (1986); Diamond (1991); Hart and Moore (1994); Petersen and Rajan (1994)).

This paper studies whether foreign banks use such mechanisms to overcome distance constraints and whether such mechanisms are effective using information on ex post loan performance. We begin by comparing the contract terms and loan pricing models of domestic and foreign banks holding constant differences in their clienteles. While previous papers have ex-

[1] Using a variety of techniques and samples, several studies document that foreign banks tend to shy away from smaller, more opaque borrowers (e.g., Berger, Klapper and Udell (2001); Clarke, Cull, Martínez Pería and Sanchez (2005); Mian (2006); Berger, Klapper, Martínez Pería and Zaidi (2008); Detragiache, Gupta and Tressel (2008); Beck and Martínez Pería (2010); De Haas, Ferreira and Taci (2010); Gormley (2010)). Others, instead, find that foreign banks are at least as likely to lend to small firms as domestic banks (e.g., Clarke, et al. (2005); Berger, Rosen and Udell (2007); Giannetti and Ongena (2009, (2012)) using transaction-based lending technologies (e.g., Berger and Udell (2006), consistent with survey evidence in De la Torre, Martínez Pería and Schmukler (2010) and Beck, Demirgüç-Kunt and Martínez Pería (2011)).

[2] Credit scoring models allow banks to compile both hard and soft information into credit scores (see, among others, Petersen (2004); Liberti and Mian (2009); Agarwal and Hauswald (2010)).

plored whether banks of different ownership lend to different firms, thereby linking different lending technologies to different firm characteristics, we eliminate the composition bias by focusing on a sample of firms that borrow from at least one domestic and one foreign bank in the same month. In particular, we compare the contract terms and loan pricing models of domestic and foreign banks when lending to exactly the *same firm* in the *same month* using both regression and matching analysis. Exploring within-firm and month variation allows us to control for unobserved firm characteristics. We are thus able to separate the two different effects of foreign bank entry—the focus on different clienteles and the use of different lending technologies and loan pricing models — by focusing completely on the latter. To investigate whether unobserved heterogeneity on the purpose of the loan drives our results, we exploit exogenous (to the firm) variation in bank ownership. Using the takeover of a domestic bank by a foreign bank, we study whether loan contract terms to the *same firm* change after the takeover using a difference-in-difference analysis.[3] In the second part of the analysis, we study the performance of domestic and foreign bank loans by looking at both the incidence of repayment problems as well as the banks' returns on loans. If domestic and foreign banks set loan conditions and price loans in an optimal way, we should not observe any systematic differences in the ex post performance of their loans, especially when holding their clienteles fixed.

Our analysis makes use of loan-level data from the public credit registry of Bolivia for the period between March 1999 and December 2003. We have detailed data on every loan such as date of origination, maturity date, contract terms, including interest rates, collateral, loan amount, and the geographical location (region) in which the loan was originated as well as data on firm characteristics such as legal structure, industry, banking relationships, and repayment behavior. The Bolivian credit market provides a good setting to analyze differences between foreign and domestic banks in their lending techniques, as it was already fully liberalized and privatized during the sample period, eliminating any distortions from the existence of government-owned banks. Moreover, foreign and domestic banks are, in principle, subject to the same regulations

[3]To control for changes in the macroeconomic conditions over the comparison period, we benchmark the changes of the treatment group relative to a control group of loans originated by all other domestic (not taken over) banks to their customers over the same period using within firm-bank variation. In additional test, we also exploit exogenous (to firm) variation in the importance of distance constraints, using a short period of socio-economic uncertainty.

such that there is no differential regulatory treatment, which could influence our results. Both the number and market shares of domestic and foreign banks are relatively balanced and stable during the sample period, enabling meaningful comparisons (Claessens and van Horen (2014)). Like for many other countries that recently opened their doors to foreign banks, Bolivia's credit markets are opaque. Many firms, for example, do not have audited financial statements, and if they do, the quality of such statements is often poor (Sirtaine, Skamnelos and Frank (2004)). It is thus important to understand how in such a setting domestic and foreign banks may be able to overcome their distance constraints and meet firms' financing needs, thus facilitating investment and growth.

Our findings suggest that foreign banks do use private contracting and credit scoring models to overcome their informational disadvantage. Such mechanisms are found to be effective in mitigating foreign banks' relative disadvantage, but as discussed below may carry important limitations. In particular, when comparing the terms of loans originated by domestic and foreign banks to the *same firm* in the *same month*, we find large and systematic differences. Consistent with theoretical predictions, foreign bank loans are, on average, 27 percentage points more likely to have collateral and have maturities that are up to 33% shorter than domestic bank loans. We also find that foreign banks charge loan interest rates that are, on average, between 89 to 107 basis points lower than the interest rates of domestic banks — a 9% discount relative to the interest rate of domestic bank loans in the sample. We obtain similar results when we examine how terms of loans to the same firms change when their (domestic) bank is taken over by a foreign bank. We find an increase in collateral by 19 percentage points, a decrease in loan maturity by around 5.4 months and a drop in loan interest rates by 58 basis points. It is thus unlikely that our results are driven by firms using domestic and foreign banks for different purposes. Additional tests distinguishing between foreign branches and subsidiaries and taking into account the time since entry or since becoming foreign owned, suggest that the observed differences in the lending practices of domestic and foreign banks are partly driven by geographical and organizational distance. In several robustness tests, we also confirm that the observed differences are not driven by borrowers playing domestic and foreign banks against each other to obtain better terms or some form of informal syndicates.

When studying how domestic and foreign banks price their loans, we find that domestic banks base their pricing on the length of their relationship with the firm, especially in the case of smaller, more opaque firms. Foreign banks instead have a more transaction-based pricing approach, relying on credit scores and collateral, especially for larger, more transparent firms. We also show that credit scores and collateral explain a larger variation in the pricing of the foreign banks loans, particularly for larger, more transparent firms. Overall, our findings indicate that domestic and foreign banks can cater to the same clientele using different lending technologies. Domestic banks seem to overcome information asymmetries in credit markets using relationship lending, while foreign banks rely more on transaction-based technologies such as asset-based lending, shorter maturities, and credit scoring models.

But are domestic and foreign banks equally successful? Do alternative lending technologies allow foreign banks to mitigate their distance constraints and profitably lend to the same firms as domestic banks? To answer this question, we study the ex post performance of domestic and foreign bank loans by looking at both the incidence of repayment problems as well as the banks' return on loans, taking into account not only the probability of default but also the loss given default. When comparing loans originated by domestic and foreign banks to the *same firm* in the *same month*, we find that loans originated by foreign banks are more likely to have repayments problems, consistent with foreign banks facing higher enforcement costs. This is more pronounced when foreign banks depart from their business model, lending without collateral and/or with longer maturities. Interestingly, we find that foreign banks are not at a disadvantage vis-à-vis domestic banks when lending with collateral and more short-term.

Our findings provide an explanation for results in Qian and Strahan (2007) and Haselmann, Pistor and Vig (2010) who find that foreign banks are more sensitive to improvements in collateral laws and creditor rights. The strengthening of formal creditor rights protection is argued to help foreign banks to more effectively use collateral and maturity to overcome cultural and informational barriers. As Haselmann, et al. (2010) put it "legal protection may offer a substitute for cultural and local knowledge." Our findings confirm this explanation. They also provide an explanations for results in Bruno and Hauswald (2014) and Claessens and van Horen (2014) who show that foreign bank entry has a positive effect in countries with more efficient

credit information sharing systems and creditor rights protection.

While we cannot make welfare statements, our findings have important repercussions for capital allocation and firms' funding structure. The different contract terms offered by domestic and foreign banks imply both a shift in the borrower population, but also in the funding structure of existing borrowers compared to a financial system with only domestic banks. It implies a shift towards firms with short-term financing needs and with pledgeable assets. Similarly, foreign bank entry could push existing clients towards more short-term investment strategies and financing of tangible assets, and possibly away from longer-term investment projects and intangible assets. On the other hand, the lower interest rates translate into lower cost of capital, thus fostering investment, in line with findings by the empirical literature on capital account liberalization (e.g., Henry (2000); Bekaert, Harvey and Lundblad (2005)).

Our paper contributes to the literature on foreign bank entry and is most closely related to Mian (2006). We both rely on credit registry data with detailed loan-level information and both investigate differences between foreign bank and domestic bank lending. While Mian (2006) focuses on differences in clienteles of foreign and domestic banks, the first effect of foreign bank entry, we explore differences in lending technologies to a given and identical clientele, the second effect of foreign bank entry. Mian (2006) finds that informational disadvantages pertaining to cultural and geographical distance constraints make foreign banks shy away from smaller, more opaque firms. We additionally find that foreign banks employ systematically different contract terms and loan pricing models even when lending to the same clients in ways consistent with foreign banks using such mechanisms to overcome distance constraints.

Our paper also relates to the broader theoretical literature on financial intermediation and incomplete contracts. The theoretical literature on financial intermediation motivates collateral as a way to mitigate adverse selection and bank losses when firms default (e.g., Bester (1985); Besanko and Thakor (1987); Bester (1987); Boot, Thakor and Udell (1991)) and ex-post frictions such as moral hazard (e.g., Boot, et al. (1991); Boot and Thakor (1994); Hart and Moore (1994); Aghion and Bolton (1997); Holmstrom and Tirole (1997)). In the absence of reliable and audited financial statements where lenders cannot rely on debt covenants to ensure entrepreneurial commitment, shorter loan maturities can also serve as an alternative disciplining

tool (e.g., Myers (1977); Berlin and Mester (1992)). As highlighted in the theoretical literature, lending at shorter maturities can help banks to better screen and monitor their clients by forcing more frequent information disclosure and renegotiation of contract terms (e.g., Barnea, Haugen and Senbet (1980); Flannery (1986); Diamond (1991); Rajan (1992); Hart and Moore (1994)). We contribute to this literature by showing that collateral and maturity can be effective mechanisms to overcome asymmetric information problems as a substitute to relationship lending.

Our paper also relates to an extensive literature on the importance of geographic and cultural distance between borrowers and lenders (e.g., Degryse and Ongena (2005); Fisman, Paravisini and Vig (2012); Beck, Behr and Madestam (2014)), which shows that loan contract terms and lending techniques are a critical function of both the geographic distance between borrowers and lenders but also the extent to which borrowers and lenders share cultural, ethnic, religious or socio-economic traits. Other papers have shown that the internal organizational distance within institutions, which can be both a function of ownership and size, matter for lending techniques (e.g., Stein (2002); Berger, Miller, Petersen, Rajan and Stein (2005); Mian (2006)). Overall, these studies find that distance increases inefficiency in credit markets. The recent and growing literature on securitization finds similar results. By creating a larger distance between the loan originator and the bearer of a loan's default risk, securitization decreases lenders' incentives to carefully screen and monitor borrowers (e.g., Keys, Mukherjee, Seru and Vig (2010); Skrastins and Vig (2014); Rajan, Seru and Vig (2015)). We contribute to this literature by showing that contract design and credit scoring can help banks overcome distance constraints.

The rest of the paper is organized as follows. Section 1.2 describes the dataset and presents descriptive statistics. Section 1.3 presents our empirical tests. Section 1.4 presents our results and several robustness checks, and Section 1.5 concludes.

1.2 Data

The paper utilizes data from the *Central de Información de Riesgos Crediticios* (CIRC), the public credit registry of Bolivia, provided by the Bolivian Superintendent of Banks and Finan-

cial Entities (SBEF). Since CIRC's creation in 1989, the SBEF requires all formal (licensed and regulated) financial institutions operating in Bolivia to record information on all loans. We have access to the entire credit registry for the period between January 1998 and December 2003. For each loan, we have information on the origination and maturity dates, the geographical location (region) in which the loan was originated, contract terms, and ex post performance. For each borrower, we have information about their legal structure, industry, bank lending relationships, and whether they have been delinquent or have defaulted on another loan in the recent past. The credit registry is used by the SBEF to monitor and supervise the banking sector. It is also used by banks to better evaluate and monitor their clients. The SBEF requires that some "hard" borrower and loan information is shared among banks to alleviate the otherwise pervasive information asymmetries in Bolivian credit market. After written authorization from a prospective customer, a lender can access the registry and obtain a report that contains information on all outstanding loans of the customer for the previous two months. Entries include the originating bank, loan amount, loan type, value of collateral, value of overdue payments, and the borrower's internal credit rating from the originating bank. Because the available information is limited to the previous two months, important information asymmetries remain. For example, if borrowers pay back an overdue loan, the record resets without any trace of overdue payments on their credit history, which borrowers may exploit strategically.[4]

The data include loans from both commercial banks and nonbank financial institutions (e.g., microfinance institutions, credit unions, mutual societies, and general deposit warehouses). To keep the set of lenders homogenous in terms of financial structure and regulation, we focus exclusively on loans granted by commercial banks to firms. Table 1.1 provides a list of the 13 commercial banks that were active in Bolivia during the sample period, seven of which are foreign-owned. Following the literature, a bank is considered foreign if at least 50 percent of its equity is owned by foreign investors (see among others Claessens and van Horen (2014)). Four of the foreign banks are branches and three are subsidiaries and, together, they account for 39% of the commercial banks' loans. Branches are integral parts of the parent company, while

[4]As shown in Ioannidou and Ongena (2010) borrowers strategically reset their credit histories before approaching new lenders. The authors also show that controlling for other observable characteristics, overdue payments on past loans — even when repayment is eventually made — are predictive of future repayment problems. Hence, this information, if made available could help banks better evaluate new customers.

subsidiaries are separate legal entities from their parent corporations. As can be observed in Table 1.1, many of the branches are part of large multinational banks with a relatively small presence in Bolivia. Most foreign banks have a lower cost of deposits than most domestic banks. On average, foreign banks pay 100 basis points lower interest rates on their deposits.

For the purposes of our analysis, we focus on commercial loans granted between March 1999 and December 2003.[5] Commercial loans represent an important segment of the credit markets for which collateral is a negotiated term that is only sometimes present. Among commercial loans, there are several types of contracts in the data, including credit cards, overdrafts, installment loans, single-payment loans, and credit lines. We focus exclusively on installment and single-payment loans and refer to these as "standard debt contracts". These contracts account for 92% of the total value of commercial loans during the sample period. Of these contracts, 98% are denominated in U.S. dollars, and we use only these loans in our analysis. To ensure the use of timely information, we only study the originations of "new loans"; renegotiations of previous loans and loans drawn on pre-existing lines of credit are excluded.[6]

All in all, this yields 32,279 loans to 2,672 firms. Table 1.2 provides summary statistics for these 32,279 loans, which we refer to as the "universe". Summary statistics are also provided separately for loans originated by foreign and domestic banks. The stars next to the mean values of domestic bank loans indicate whether the differences between domestic and foreign banks are statistically significant. Throughout the text, ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively. Table A.1 in the Appendix provides detailed definitions for all variables used.

As can be observed in Table 1.2, 47% of loans are installment loans. The average loan amount is US$162,000, 25% of them are secured, and the average loan maturity is 11 months. In terms of price, the average loan carries an interest rate of 13.5% and a spread of 10%

[5] Although we have data as of January 1998, we start our sample in March 1999 since prior to this date the data do not allow us to distinguish between commercial and consumer loans. We use prior information from January 1998 through February 1999 to fill in the history of bank-firm relationships as well as the firms' credit history.

[6] Renegotiations are identified as follows. Banks are required to indicate whether a new loan is a renegotiation of a previous performing or nonperforming loan. We use this information to exclude renegotiation. Loans drawn on pre-existing lines of credit, instead, are identified as follows. When a borrower draws on a pre-existing line of credit, a "new loan" appears in the registry with an origination date and contact terms as of the date the bank originated the credit line. Since the date the loan first appears in the registry is subsequent to the origination date, we can identify when a "new loan" is a draw on a pre-existing line of credit and exclude it from our sample.

over the rates of US Treasury Bills of comparable maturity. In terms of firm characteristics, 49% of loans are given to limited liability firms. The remaining loans are to firms where either all or some partners have unlimited liability. This includes joint stock companies (23%), limited partnerships (13%), sole proprietorships (13%, and general partnerships (1%). Table A.1 provides for definitions of each legal structure. The average firm in the sample maintains an outstanding debt of around US$2,000,000. Only 21% of loans are given to firms with recent repayment problems (i.e., delinquency or default in the past year). At origination, around 87% of loans have the best rating (1), 10% have the second best rating (2), and 3% have the two worst ratings (3 and 4). Around 56% of loans are given to firms with multiple bank-lending relationships. The average relationship length is 22 months, 26% of loans are given to firms with additional lending products at the bank, and 71% of loans are given to firms with at least 50% of their outstanding loans are from the bank.

When comparing the terms of domestic and foreign bank loans, some striking differences emerge. Loans originated by foreign banks are on average larger by around US$70,000 and carry interest rates that are lower by around 66 basis points. Foreign bank loans are also more likely to be secured: 38% of the foreign bank loans have collateral, while only 16% of domestic bank loans do. Their maturities are also shorter by around 4 months. While these differences are consistent with foreign banks employing different lending technologies and offering different contract terms, they could also be explained by differences in firm composition as domestic and foreign banks may favor or attract different types of firms. This highlights the importance of holding constant differences in their clienteles.

In fact, as can be observed in Table 1.2, the clients of foreign banks are, on average, larger and riskier firms with "weaker" bank-lending relationships. In particular, loans originated by foreign banks are less likely to be given to sole proprietorships and more likely to be given to joint stock companies, which are typically larger firms. The firms' outstanding bank debt is also substantially larger among foreign bank loans (by around US$680,000), consistent with a distribution tilted towards larger firms. When looking at credit quality, we also observe that the incidence of recent repayment problems is substantially higher among foreign bank loans and that credit ratings are worse, suggesting that foreign banks may have riskier pool

of firms and/or higher enforcement costs. The incidence of multiple relationships is higher among foreign bank loans and the average relationship length is shorter by around 2 months, consistent with shorter and more dispersed relationships. Other products from a bank and a "primary bank" status are also less likely among foreign bank loans, consistent with overall weaker relationships.

To understand whether differences in contract terms between domestic and foreign bank loans are solely due to their different clienteles or also due to the use of different lending technologies, we eliminate the firm-composition effect by comparing the contract terms of domestic and foreign bank loans to the same firm in the same month. To this end, we restrict our analysis to a sub-sample of loans to firms that received a new loan from at least one foreign and one domestic bank in the same month. (In robustness tests we also confirm our results for the excluded sample.) The restriction results in a sub-sample of 5,137 loans to 287 firms. This sub-sample constitutes 25% of the total lending amount of the entire sample. The second part of Table 1.2 provides summary statistics for our sub-sample and compares it to the "universe" of all loans. Like in the first part, statistics are provided for all loans in the sample as well as for foreign and domestic banks separately. The stars next to each mean value indicate whether it is statistically different from its corresponding value for the entire sample.[7]

As can be observed in Table 1.2, our sub-sample draws more heavily on the largest firms. This is true for both domestic and foreign banks. For example, the average loan amount and the outstanding bank debt are significantly higher for both groups. Similarly, sole proprietorships are much less common, while joint stock companies are more frequent. With respect to credit quality results are somewhat mixed. The incidence of past repayment problems is higher in our sub-sample, suggesting riskier firms. This is true for both domestic and foreign banks. However, when looking at credit ratings the picture is different. While the ratings for foreign bank loans are worse, the ratings of the domestic bank loans are better, despite their worse credit histories. It is hard to know what drives this discrepancy. It is possible, for example, that domestic banks are inflating their ratings to relax regulatory constraints such as loan

[7]In particular, we construct a specific test statistic with a correspondingly adjusted variance for our mean comparisons. This test statistic accounts for the fact that we compare the means of the entire sample with the means of a sub-sample. The derivation of the statistic and its asymptotic behavior is available upon request.

loss reserve requirements. With respect to relationship characteristics, we find that firms in our sample have on average longer relationships with their banks but are less likely to have a primary bank, which is expected given that we focus on firms with multiple relationships that tend to be larger.

In regression analysis, we also present results for subsamples of smaller and larger firms within the sample of firms that borrow from both domestic and foreign banks in the same month. As information on total assets is not available (for confidentially reasons the firm identities were altered before releasing the data to us), we approximate firm size with total bank debt.[8] For each firm we calculate the average outstanding debt across all financial institutions over the sample period and divide firms into smaller and larger firms using the median firm's total bank debt (US$1,014,978). In robustness tests, we also try alternative sample splits using instead the 30th and 70th percentiles of the total bank debt (US$466,568 and US$2,329,930, respectively).

1.3 Methodology

We examine whether domestic and foreign banks employ different lending technologies by conducting two sets of empirical specifications. We first test whether loans originated by domestic and foreign banks have systematically different loan contract terms. We compare the incidence of collateral, the maturities, and the interest rates of loans originated by domestic and foreign banks to the same firm in the same month, controlling for several other factors that might explain any observed differences between them. In a second step, we also explore whether domestic and foreign banks use systematically different factors to price their loans. Specifically, we gauge whether the incidence of collateral, the rating of the firm by the bank and the length, depth and breadth of the relationship affect the pricing of loans and whether these relationships vary significantly across banks of different ownership.

[8]Using data from the U.S. Survey of Small Business Finances (SSBF), Petersen and Rajan (1994) find that conditional on the existence of institutional debt, the ratio of bank debt to total debt does not vary significantly with firm size. For firms with less than $15,000 in total assets, this ratio is equal to 51 percent, while for firms with more than $2,000,000 in total assets the ratio equals 62 percent. Since capital markets are less developed in Bolivia than in the US, this ratio may vary even less, such that total bank debt is most likely highly correlated with firm size.

To investigate whether domestic and foreign banks impose systematically different loan conditions on their clients we estimate the following model using Ordinary Least Squares (OLS):

$$LoanContract_{ijkt} = \alpha_1 + \beta_1 ForeignBank_k + \beta_2 Firm_{jkt} + \beta_3 Loan_{ijkt} + \eta_j \times \gamma_t + \varepsilon_{ijkt}, \quad (1.3.1)$$

where i, j, k, t index loans, firms, banks, and time (month-year) of loan origination, respectively. For $LoanContract_{ijkt}$ we employ three dependent variables: (i) a dummy indicating whether the loan contract includes the pledge of collateral, (ii) the natural logarithm of one plus the number of months between loan origination and maturity, and (iii) the loan interest rate minus the rate on US Treasury securities of comparable maturity at loan origination. Our key explanatory variable, $ForeignBank_k$, is a dummy variable that indicates whether the originating bank is foreign-owned. This variable is time invariant as during the sample period there are no changes in bank ownership. In a robustness test, however, we extend our sample period backwards to study the take over a domestic bank by foreign investors.

Our set of control variables includes several indicators that control for firm and other loan characteristics as well as firm-month fixed effects. The vector $Firm_{jkt}$ is comprised of firm characteristics that vary within the same month across banks. This includes the firm's internal rating at each bank as well as indicators of the strength of the bank-firm lending relationship at the time of the loan origination. For ratings, we include two dummy variables: $Rating2_{jkt}$ and $Rating3\&4_{jkt}$. $Rating2_{jkt}$ equals 1 if the firm's rating equals 2, and equals zero otherwise, while $Rating3\&4_{jkt}$ equals 1 if the firm's rating equals either 3 or 4, and equals zero otherwise.[9] ($Rating1_{jkt}$, which indicates the best rating, is the omitted category.) These ratings refer both to "hard" information about firms as well as "hardened" soft or subjective information collected by the loan officer in the screening process. They can vary across banks for different reasons, including different soft and hard information sets and different interpretation of available information. To gauge the intensity of a bank-firm relationship we employ three variables: $RelDuration_{jkt}$, $RelScope_{jkt}$, and $PrimaryBank_{jkt}$. $RelDuration_{jkt}$ is equal to the natural logarithm of one plus the number of months we observe the bank-firm pair in

[9]To gain statistical power, ratings 3 and 4 are merged into one dummy variable given the small number of loans with such ratings (see descriptive statistics in Table 1.2).

a credit relationship;[10] $RelScope_{jkt}$ is a dummy variable indicating whether the firm has other credit products from the bank (e.g., credit cards, overdrafts, mortgages); and $PrimaryBank_{jkt}$ indicates whether the bank accounts for more than 50% of the firm's bank debt.

The vector $Loan_{ijkt}$ includes other loan characteristics. $Installment_{ijkt}$ is a dummy variable indicating whether the loan is an installment loan as opposed to a single-payment loan and $LoanAmount_{ijkt}$ is measured as the natural logarithm of one plus the amount of loan proceeds at origination in US dollars. In the interest rate regressions, we also include $Collateral_{ijkt}$, and $Maturity_{ijkt}$, effectively assuming that these contract terms are determined prior to the loan interest rate, consistent with the maintained assumption in most of the extant empirical literature in banking (e.g., Berger and Udell (1995); Elsas and Krahnen (1998); Harhoff and Körting (1998); Degryse and Ongena (2005); Ortiz-Molina and Penas (2008)). Since each of these other contract terms may be simultaneously determined with the dependent variable and is thus potentially endogenous, we first estimate our models without other contract terms among the controls. In robustness tests, we also employ a matching technique similar to Ioannidou and Ongena (2010). By matching on other contract terms we do not need to assume anything about the decision process. Matching is nonparametric (imposing no functional form restrictions on the way the matching variables relate to the dependent variable) and does not incorporate any information from outside the overlap region between the foreign bank loans and the domestic loans.[11] We also provide additional robustness tests about the potential role of other omitted variables.

Finally, $\eta_j \times \gamma_t$, are firm fixed effects, η_j, interacted with time (month-year) fixed effects, γ_t, to account for observable and unobservable firm characteristics. Our estimates are thus obtained using only within firm-month variation for the sub-sample of firms that borrow from both domestic and foreign banks in the same month. This is the reason for which we do not include time-invariant firm characteristics. In addition to estimating the above regression for the whole sample, we also estimate it separately for firms of different size using their total

[10] $RelDuration\,Square_{jkt}$ (i.e., the square of $RelDuration_{jkt}$) is also sometimes included in our specifications. For our baseline results, we estimated the regressions with and without the square term to test for the possibility of a non-linear relationship. If the square term is found to be statistically significant, we report the regression results with the square term. If not, we report the results without the square term.

[11] See, for example, discussions in Ioannidou and Ongena (2010) and Roberts and Whited (2012).

outstanding bank debt as a proxy of their size. We use OLS for all three dependent variables. For collateral, we use OLS instead of a non-linear Probit or Logit model as we would otherwise lose a large number of firm-month clusters with no variation in collateral across different banks for the same firm within the same month.[12] In all cases, the standard errors are clustered at the firm level to account for the possibility that the observations of the same firm across different loans, bank, and time are correlated with each other.[13]

Given our set of controls variables, a statistically significant β_1 would indicate that the probability of collateral, the maturity and the interest rates on domestic and foreign bank loans to the same firm in the same month are systematically different, even after controlling for possible differences in banks' credit risk assessment, the strength of the bank-firm lending relationship, as well as other loan contract terms. If foreign banks are using collateral and maturity to mitigate informational and enforcement disadvantages, we would expect that they will be more likely to require collateral and lend at shorter maturities than domestics banks, even when lending to the same customers.

Next, we examine whether the factors that explain the variation in the loan interest rates to the same firm in the same month vary systematically between domestic and foreign banks by introducing interactions between the foreign bank dummy and our control variables:

$$\begin{aligned} LoanSpread_{ijkt} = \alpha_2 + \gamma_1 ForeignBank_k + \gamma_2 Firm_{jkt} + \gamma_3 Loan_{ijkt} + \gamma_4 ForeignBank_k \times Firm_{jkt} \\ + \gamma_5 ForeignBank_k \times Loan_{ijkt} + \eta_j \times \gamma_t + \varepsilon_{ijkt}, \end{aligned} \tag{1.3.2}$$

where i, j, k, t index loans, firms, banks, and time (month-year), respectively. All variables are defined as in equation (1.3.1). In this case, our focus is on the coefficients of the interaction terms, which indicate whether internal credit ratings, relationship strength, and other loan

[12] Because of the large number of fixed effects in our model relative to the smaller number of periods for which a borrower is observed, a non-linear model could also give inconsistent estimates; this is known as the "incidental parameter problem" (see, for example, discussion in Cameron and Trivedi (2005, pp. 726-727).)

[13] In unreported sensitivity analyses, we also test the robustness of our results to alternative clustering assumptions. Standard errors may be correlated at the bank-level if bank-level shocks affect all loans given by a specific bank. We confirm our results when clustering standard errors at both the bank-level and at the bank-year and firm level.

contract terms are used differently by domestic and foreign banks when pricing their loans. Since the coefficients are again estimated using only within firm-month variation such differences would point to the use of different lending technologies. If, for example, foreign banks rely more on transaction-based technologies, such as credit scoring and asset-based finance, we would expect credit ratings and collateral to play a more prominent role in foreign banks' pricing. Similarly, if domestic banks rely more on relationship lending, we expect the relationship variables to be more important for the domestic banks' pricing. As in the case of regression (1.3.1), the model is estimated using OLS clustering the error terms at the firm-level.

1.4 Results

1.4.1 Bank Ownership and Loan Contract Terms

The results in Table 1.3 show that foreign bank loans are more likely to have collateral and are more short-term than domestic bank loans. Columns I to V present results of the collateral regressions and Columns VI to X present results of the maturity regressions. As discussed earlier, both the collateral and the maturity regressions are estimated using OLS. In all cases, we include firm×time fixed effects and investigate whether the likelihood of pledging collateral and the maturity of loans vary systematically between domestic and foreign banks, even when lending to the same firm in the same month.

As can be observed in Columns I to III, foreign bank loans are between 27 and 28 percentage points more likely to have collateral; a large effect given that only 33% of all loan contracts in our sample include collateral. We also find that loans to firms with a rating of 2 rather than 1 are more likely to have collateral. Loans to firms with longer relationships and a primary bank status with a bank are less likely to have collateral.[14] These results support previous empirical findings in the relationship lending literature (e.g., Petersen and Rajan (1994); Berger and Udell (1995); Elsas and Krahnen (1998); Harhoff and Körting (1998); Machauer and Weber (1998)). Controlling for other contract terms has no material effect on our findings. The coefficients of the other contract terms indicate that larger loans are more likely to have collateral. In

[14] Adding the square of $RelDuration_{jkt}$ in the specification results in statistically insignificant coefficients for $RelDuration_{jkt}$ and its square term.

Columns IV and V, we also split the sample into smaller and larger firms and confirm the results for both groups. The results in Columns VI-VIII show that loans granted by foreign banks have maturities that are between 19% and 27% shorter than loans granted by domestic banks. At the average maturity of nine months, this implies a difference of two to three months. With respect to our control variables, we find that variation in credit ratings is not significantly associated with variation in loan maturities. Relationship characteristics, on the other hand, seem to play an important role. Loans to firms with longer relationships (more than eleven months) and a primary bank status with a bank have longer maturities, consistent with the relationship lending literature. Additional products, on the other hand, are associated with shorter maturities, although this effect is significant only at the 10% level. Controlling for other contract terms results in a larger absolute coefficient for $Foreign$, suggesting that loan maturity is correlated with $Installment$ and $LoanAmount$. The coefficients of the other contract terms indicate that installment loans and loans with larger loan amounts have longer maturities. Dividing our sample into smaller and larger firms in Columns IX and X, confirms again our findings for both sets of firms.[15]

The results in Table 1.4 show that foreign banks charge the same firm in the same month about 89 to 107 basis points lower interest rates than domestic banks, confirming the difference reported in the descriptive statistics. Relative to the interest rate of domestic bank loans, these estimates imply a 9% discount. As can be observed in Column I, loans originated by foreign banks carry on average 89 basis points lower interest rates than loans originated by domestic banks. This difference is statistically significant at the 1% level and remains unchanged when we additionally control for possible differences in credit ratings and the strength of the bank-firm relationships in Column II. Adding other contract terms in Column III reveals that collateralized loans and loans with longer maturities have on average lower interest rates. The negative coefficient of collateral is consistent with the theoretical literature on the role of collateral in debt

[15]We also explore whether collateral and maturity are complements or substitutes as mechanisms to overcome information asymmetries and agency problems. We split domestic and foreign bank loans within our sample into four groups according to whether (i) they are secured or unsecured and (ii) have a maturity below or above the median maturity in our sample. For foreign bank loans, we find that 65% of their secured loans have a below-median maturity, thus two-thirds of secured foreign bank loans also have a short maturity. On the other hand, among unsecured foreign banks loans, the distribution between below-median (52%) and above-median (48%) loans is much more even. There is no comparable difference for domestic bank loans. In summary, foreign banks seem to use collateral and short maturity as complements rather than substitutes.

contracts.[16] The negative coefficient on maturity is consistent with lower risk firms (projects) obtaining loans with longer maturities. Controlling for other contract terms in Column III does not qualitatively change the foreign bank result, but the estimated coefficient becomes bigger in absolute terms: it increases from -89 to -107 basis points, suggesting that loan interest rates are correlated with other contract terms and bank ownership. With respect to other control variables, we find that lower credit ratings are associated with significantly higher interest rates, while stronger lending relationships are typically associated with lower interest rates. Dividing the sample into smaller and larger firms confirms again our findings for both groups (Columns IV and V). The estimated difference is 113 basis points for larger firms and 96 basis points for smaller firms.[17] The remaining columns of Table 1.4 show that differences in market shares and funding costs do not explain the interest rate differential.[18] $ForeignBank$ continues to enter with the same economic and statistical significance as before, while neither funding costs nor market shares enter significantly.

Overall, our findings suggest that foreign banks are more likely to require collateral, grant loans with shorter maturities, and charge lower interest rates, even when lending to the same firm in the same month. The foreign bank discount seems to be over and above of what is implied by the other contract terms they offer and is not explained by funding costs or market power. There is thus a clear trade-off for firms when taking out loans from both domestic and foreign banks, which may explain why they maintain active relationships with both domestic and foreign banks. There could be, however, concerns about sample selection, omitted variable

[16]The lower interest rates on secured loans are consistent with both the ex-ante and the ex post theories of collateral. Under the ex ante theories unobservably safer borrowers are more likely to pledge collateral to signal their quality and receive lower interest rates (e.g., Bester (1985, 1987)). Under the ex post theories, observably riskier borrowers are more likely to be required to pledge collateral to mitigates ex post frictions such as moral hazard and bank losses in the default (e.g., Boot, et al. (1991); Boot and Thakor (1994); Holmstrom and Tirole (1997)). Hence, conditioning on borrower risk, the ex post theories are also consistent with lower interest rates for secured loans.

[17]Similar results are obtained if we instead split firms using the 30th and 70th percentiles of the total outstanding bank debt with a 113 basis point estimated difference for larger firms and 92 basis points for smaller firms. The somewhat larger discounts for larger firms may be capturing economies of scales and smaller average costs of screening and monitoring larger loans and firms. This would be consistent with the presence of important fixed costs, for example.

[18]Table 1.1 shows that foreign banks have significantly lower funding costs and many of them have smaller market shares than domestic banks. These differences can in principle drive the interest rate differential that we find. Lower funding costs can allow foreign banks to charge lower loan interest rates. Higher market shares may be associated with higher or lower interest rates depending on whether market power or economies of scale effects dominate.

biases, and concerns about the functional form in which contract terms relate to each other. Next, we subject our results to several robustness tests to address these concerns.

1.4.2 Robustness Checks

First, we re-estimate equation (1.3.1) on a sample of firms that did *not* have loans from both domestic and foreign banks in the same month. We thus use the 27,142 observations from the universe of observations presented in Table 1.2 that were not part of the regression sample so far (i.e., firms that borrow either from domestic or from foreign banks but not both). Different from the previous regressions, we therefore include firm- and month-fixed effects, but not their interaction. While this sample does not allow for the clean identification that we have used so far, it shows whether the systematic differences in loan contract terms between domestic and foreign banks hold for the larger population of firms in Bolivia. As can be observed in Table 1.5, we find that the differences between domestic and foreign banks for this larger sample of firms are qualitatively similar to those reported for our restricted sample. The economic size of the effects is smaller for the interest rate spread and collateral and larger for the maturity difference. Overall, these results confirm that our findings are relevant beyond our limited and selective sample of enterprises with access to both domestic and foreign banks at the same time.[19]

To gauge the relative importance of possible omitted variable bias, we develop a test in the spirit of Altonji, Elder and Taber (2005) following Bellows and Miguel (2009). The test compares the coefficient estimates from a univariate model as in Columns I and VI of Table 1.3 for collateral and maturity and in Column I of Table 1.4 for the interest rate with coefficient estimates including a full set of control variables (Columns III and VIII of Table 1.3 and Column III of Table 1.4, respectively). In all cases, the test statistics indicates that the covariance of the unobserved factors with bank ownership would have to be a multiple of at least 3.4 times of the covariance of the observed control variables with bank ownership, indicating that it is

[19]In unreported robustness tests, we also split the sample into smaller and larger firms as proxied by total loan amount outstanding. As in Tables 1.3 and 1.4, we find a higher incidence of foreign banks asking for collateral among smaller than larger firms, but less of a maturity discount for foreign bank loans among small firms and less of an interest rate differential. This suggests that collateral is a contract feature attractive for foreign banks to overcome information asymmetries with smaller firms, while maturity is a contract feature attractive more for larger firms.

unlikely that results are driven by omitted variable biases and selection on unobservables.[20]

We also investigate the robustness of our findings using a matching technique instead of a regression analysis. Matching is nonparametric and thus imposes no functional form restrictions on the way the matching variables relate to the dependent variable. Moreover, as it uses only observations that satisfy the matching criteria, it does not incorporate any information from outside the overlap region, mitigating omitted variable concerns from the joint determination of contract terms. We match using a procedure similar to Ioannidou and Ongena (2010) and Degryse, Ioannidou, and von Schedvin (2015) matching on each variable individually allowing for replacement and multiple neighbors. For discrete variables, we use exact matching. For continuous variables, we use caliber matching using 0.5 a standard deviation radius for each of our matching variables. Table 1.6 reports our findings. In Column I, we match on firm identity and month of loan origination. In Column II, we additionally match on ratings and relationship characteristics. In Column III, we also match on other contract terms as in Tables 1.3 and 1.4. We present results that correspond to our specifications in Tables 1.3 and 1.4. As can be observed in Table 1.6, in all cases we find results that are qualitatively similar to those presented earlier.

Next, we also examine whether unobserved differences on the purpose of the loan are driving our results by exploiting exogenous (to the firm) variation in distance constraints. For the first test, we exploit exogenous (to the firm) variation in bank ownership using the take-over of the domestic Banco Boliviano Americano by the foreign Banco de Crédito de Bolivia in May 1999— the only take-over event in our sample period. The take-over of a domestic bank by a foreign bank is expected to increase distance constraints as it increases the geographical distance between the bank's headquarters and the local branches. It may also increase organizational distance by adding more layers in the credit allocation decision, favoring the use of more

[20]The test measures how large the selection on unobservables relative to the selection on observables needs to be to account for the entire foreign bank effect under the null hypothesis of no average treatment effect. Specifically, we calculate the ratio between the value of the coefficient in the regression including controls (numerator) and the difference between the coefficient derived from a regression without covariates and the coefficient from the regression including controls (denominator). This ratio shows how strong the covariance between the unobserved factors explaining loan conditionality and bank ownership needs to be, relative to the covariance between observable factors and bank ownership, to explain away the entire effect we find. We find ratios of 23.1, 3.4, and 5.9 for the collateral, maturity, and interest rate regressions, respectively, suggesting that it is rather unlikely that our findings are simply driven by omitted variables.

standardized process in evaluating and managing risk such as collateral and maturity. To perform this test, we identify firms that between January 1998 and the month prior to the merger received new loans from their initial bank (Banco Boliviano Americano).[21] We match these loans with any loans that the same firm received from the acquiring bank in the two years after the acquisition, dropping firms that prior to the acquisition were already customers of the acquiring bank (Banco de Crédito de Bolivia). This gives us 85 loans and 140 pairs. To control for other changes in the macroeconomic environment and the banking system over the comparison period, we use as a control group, loans to customers of all other domestic (not taken over) banks before the acquisition and loans to the same customers after the acquisition. Changes for the control group are calculated after matching on both firm and bank level (i.e., we compare within firm and bank changes in the contract terms). Results are reported in Table 1.7. We find a large and significant increase in the likelihood of collateral for the treatment group as well as a drop in maturity and interest rates after their bank has been acquired by a foreign bank. Relative to the control group, the likelihood of collateral for the treatment group increased by 18.8 percentage points, maturity decreased by around 5.4 months, and spreads decreased by 36 basis points, respectively. Overall, the results show that the exogenous to the firm change in bank ownership had similar effects on loan conditionality for a given group of firms as those documented earlier.

All in all, these results confirm that foreign banks are more likely to require collateral, grant loans with shorter maturities, and charge lower interest rates, even when lending to the same firm in the same month. These results are robust to controlling for sample selection, possible omitted biases and do not seem to be driven by firms using foreign and domestic bank loans for different purposes.

[21] We extend the sample period backwards for this exercise to January 1998. A disadvantage of using data prior to 1999 is that we cannot distinguish between commercial and consumer credit. Matching on firms and using a difference-in-difference analysis allows us to mitigate this problem. At the beginning of the sample period, one more bank changed ownership status. In particular, Banco Solidario's ownership changed from domestic to foreign owned when one of its three main investors moved outside Bolivia, making this a less clear case.

1.4.3 Bank Ownership and Loan Pricing

We now turn to the second empirical model to gauge differences between foreign and domestic banks in their loan pricing. In particular, Table 1.8 reports results for the fully interacted model of equation (1.3.2). Column I shows the estimated coefficients for domestic banks, Column II reports the coefficients of the interaction terms with the foreign bank dummy (i.e., the difference of foreign banks relative to domestic banks), and Column III shows the cumulative coefficients for foreign banks. In Column IV-VI and VII-IX, we also report the corresponding specifications for smaller and larger firms, respectively.[22]

The results in Table 1.8 show significant differences between domestic and foreign banks in the pricing of their loans even when lending to the same firm in the same month. The results in Columns I to III reveal that only foreign banks use credit ratings to price their loans. Specifically, we find that the variation in credit ratings is significantly related to the variation in interest rates in the case of foreign but not in the case of domestic banks.[23] Moreover, as can be observed in Columns VI and IX, foreign banks use credit ratings mainly for the pricing of their loans to larger firms.[24] Domestic banks instead seem to base their pricing on the strength of their lending relationship with the firm, particularly for smaller firms. As can be observed in Column IV, smaller firms with longer relationships above seven months and a primary bank status with a domestic bank are charged lower interest rates.[25] These effects are less pronounced for larger firms. While the coefficient of the primary bank status in Column

[22]As before, smaller and larger firms are defined using the median firm's total outstanding bank debt as a threshold as in Table 1.3. However, similar results are obtained if we use the 30th and the 70th percentiles instead.

[23]As this finding might be due to the lower variation in credit ratings by domestic banks, we re-ran these regressions, using a standardized credit rating variable, where each rating is expressed as the difference between the actual rating and the originating bank's average rating divided by the standard deviation in the originating bank's rating. The results using these standardized credit ratings, available on request, confirm our findings.

[24]In unreported results, we re-estimate our specifications of equation (1.3.2) after replacing the credit ratings with a dummy variable indicating whether the firm had observable defaults or repayment problems in the past year following Berger, et al. (2011).This variable is not found to be statistically significant, suggesting that credit ratings contain additional (perhaps more forward looking) information about the firm than mere past (non) performance. This might be explained by the fact that by looking at a sample of firms that get multiple loans from multiple banks in a given month we focus on some of the larger firms in the sample. For example, re-estimating the interest rate regression for the entire sample yields a positive and significant coefficient for the past nonperformance variable.

[25]The positive coefficient on relationship length and the negative coefficient on its square imply that interest rates decrease in relationship length, when Rel Duration is larger than 2.08, corresponding to a relationship length of 7 months. The negative relationship between relationship length and interest rates turns significant at 26 months.

VII remains statistically significant, the size of the coefficient is much smaller (in absolute terms) and relationships length is not found to matter for larger firms. With the exception of additional products from the bank (*Rel Scope*), which are positively related to interest rates, none of the relationship characteristics is found to explain the interest rate variation of foreign bank loans to larger firms in Columns VI and IX. Turning to other contract terms, we observe that collateral pledges are associated with lower interest rates for larger firms in the case of foreign banks, but not in the case of domestic banks. Installment loans are charged higher interest rates by both domestic and foreign banks, but only in the case of larger firms. While variation in the loan amount is not significantly associated with variation in interest rates, higher maturity loans attract lower interest rates, both from domestic and foreign banks.

Table 1.9 provides additional insights on the extent to which hard or hardened information explain variation in the pricing of foreign and domestic bank loans. Following Rajan, et al. (2015), we estimate a separate pricing model for foreign and domestic bank loans using credit ratings and collateral as the only explanatory variables and compare the resulting *R-squares* for domestic and foreign bank loans. If foreign banks rely mostly on hard or hardened information while domestic banks rely in addition on soft information stemming from the depth and breadth of the relationship with borrowers, then credit ratings and collateral should explain a higher share of interest rates in the case of foreign than in the case of domestic banks. We also estimate this model separately for smaller and larger firms. As can be observed in Columns I and II of Table 1.9, ratings and collateral explain about 10% of the interest rate variation of foreign bank loans and a mere 2% of the domestic bank loans. Re-estimating these models for smaller and larger firms separately confirms these results for both smaller and larger firms and shows that the difference in *R-squares* between foreign and domestic banks is more pronounced for larger firms. For smaller firms, the foreign banks' *R-square* is about 76 percent higher than that of domestic banks, while it is 574 percent higher for larger firms. Overall, these findings suggest that hard or hardened information such as credit scores and collateral play a much more important role in the pricing of foreign bank loans, particularly for larger firms. This is not surprising as larger firms are more likely to have collateral and more information may be available on them allowing a more meaningful use of credit scoring models. Nevertheless, as can

be observed in Table 1.9 significant heterogeneity remains consistent with the extant literature on the pricing of commercial loans (see, for example, Cerqueiro, Degryse and Ongena (2011)).

All in all, these findings are consistent with foreign banks using transaction-based technologies, such as credit scoring and asset-based finance that rely on hard or hardened information, especially when lending to larger firms. Domestic banks instead engage more in relationship lending, especially in the case of smaller firms. These results confirm the new paradigm that domestic and foreign banks cater to the same clientele using different lending technologies (e.g., Berger and Udell (2006); De la Torre, et al. (2010)). However, it also highlights the need to adequately control for differences in clientele as the use of different lending technologies for firms of different size could produce similar but misleading results.

1.4.4 What Drives "Distance Constraints"?

In this section, we present additional results exploiting cross-sectional variation on "distance" using the legal form of the foreign bank, time since entry or becoming foreign owned, and geographical and organizational distance between borrowers and lenders. These results could shed light on the source of distance constraints between domestic and foreign banks.

First, we explore differences between two different legal forms of foreign bank presence: foreign branches and foreign subsidiaries. During the sample period, Bolivia had four foreign branches and three foreign subsidiaries. Branches are integral parts of their parent company, while subsidiaries are separate legal entities. Hence, while in the case of subsidiaries most lending decisions will be delegated to local management, in the case of branches, it is likely that there will be less local discretion and more centralization resulting in greater geographical and organizational distance, resulting in the use of different lending technologies with greater emphasis on hard information (e.g., Aghion and Tirole (1997); Stein (2002); Mian (2006); Skrastins and Vig (2014)). In addition, subsidiaries are former domestic banks that may have better knowledge of the country (local economy, corporate culture, legal and institutional framework) than branches of foreign multinational banks.[26] We thus expect that the differences between domestic and

[26]The literature also suggests that branches and subsidiaries follow different business models. As discussed in Cerutti, Dell'Ariccia and Martínez Pería (2007) foreign branches are smaller operations focusing on small segments of the overall market such as wholesale operations and investment banking, with less focus on retail operations.

foreign banks in their lending technologies may be even more pronounced for foreign branches than for foreign subsidiaries. Hence, in Panel A of Table 1.10 we re-estimate a slightly modified version of equation (1.3.1), in which the foreign bank dummy is split into its two components and the sample is restricted to firms that obtained at least one loan from a domestic bank, a foreign subsidiary, and a foreign branch in the same month. The resulting sample includes 689 loans to 30 unique firms.

As can be observed in Panel A, both foreign branches and foreign subsidiaries charge lower interest rates and demand more collateral than domestic banks and these differences are more pronounced for the foreign branches. With respect to maturity, we find that while foreign subsidiaries have shorter maturity loans relative to domestic banks, foreign branches do not. In Panel B we confirm these results using a larger sample. In particular, we relax the restriction that firms must be borrowing from all three types of banks in the same month and consider a sample of firms borrowing from the three types of banks at any point during the sample period, resulting in a sample of 7,040 loans to 117 firms. We thus replace the firm-month fixed effects with firm and time fixed effects. Results are similar to those presented in Panel A. Overall, the results in Panels A and B of Table 1.10 confirm our findings for both types of foreign banks and give additional insights into the different disciplining and risk mitigation mechanisms chosen by foreign branches and foreign subsidiaries. While foreign branches focus on having collateral for their loans, foreign subsidiaries focus more on shorter maturity loans.

Next, we examine how the foreign banks' may vary with time since entry or since becoming foreign owned. We distinguish foreign banks between branches and subsidiaries as the impact of time on these two groups may be different. Branches may become more like domestic banks as time passes and they become more accustomed with the local economy, its culture, and institutions. Subsidiaries are former domestic banks that have been taken over by foreign investors. Subsidiaries may thus become more like foreign banks as time passes and begin adopting the process and policies of their parent company. To investigate this possibility, we re-estimate equation (1.3.1) allowing for interactions with time since entry or time since becoming foreign owned, defined as the natural logarithm of the number of years between the date of loan origination and the date the bank entered Bolivia or became foreign owned. We

estimate this specification using the larger sample in Panel B as we do not have sufficient variation for the restricted sample of 689 loans of Panel A. Results are presented in Panel C of Table 1.10.

We find that as time since entry increases, foreign branches become more like domestic banks. In particular, while foreign branches are more likely to require collateral and lend at shorter maturities than domestic banks, these differences become less pronounced as time since entry increases. In terms of economic significance, starting from the mean time since becoming foreign owned, a one standard-deviation increase in time since entry (by 1.85 years) is associated with a 6 percentage points decrease in the relative likelihood of collateral and a 52 percent decrease in the maturity differential. Foreign subsidiaries instead begin to behave more like foreign banks as time since becoming foreign owned increases. They begin to require more collateral and lend at shorter maturities as the time passes. In terms of economic significance, starting from the mean of around 5 years, a one standard deviation increase in the time since foreign owned (by 2.17 years), is associated with a 13 percentage points increase in the relative likelihood of collateral and a 50 percent increase in the maturity differential by around 2 months. In terms of loan interest rates, we find that in both cases the foreign bank discounts increase with time. In terms of economic significance, a one standard deviation increase in time since entry or since becoming foreign owned is associated with a further widening of the interest rate discount by 22 basis points for foreign branches and 35 basis points for foreign subsidiaries. Interestingly, the direction as well as the economic magnitudes of the effects of time since becoming foreign owned are similar to those documented in Table 1.7 for the take-over of the former domestic bank from foreign investors.

Next, we study whether the distance between the region of loan origination and a bank's headquarters drive part of the observed differences between domestic and foreign banks. We re-estimate the specifications in Panel B allowing for interactions with *Non-Local*, a dummy variable that indicates whether the bank's headquarters are in a different location (region or country) than where the loan is originated.[27] This variable intends to capture greater

[27]While we know the precise location of a bank's headquarters, we do not know the precise location of the branch that originates each loan or the precise location of each firm. We can only observe the region in which each loan is originated. The data availability does not allow us to further study, for example, differences due to special price discrimination emanating from transportation costs and market power as in Degryse and Ongena

geographical and organizational distance between loan origination and the bank's headquarters. The estimation results are presented in Panel D. The omitted group in these specifications is local domestic banks.[28] Hence, the coefficients of *Foreign Branch* measure differences of non-local foreign branches vis-à-vis local domestic banks. The coefficients of *Foreign Subsidiaries* measure the differences between local foreign subsidiaries and local domestic banks, while the combined coefficients (reported at the bottom of the table), measure differences between non-local domestic *Foreign Subsidiaries* and local domestic banks. The coefficients of *Non-Local* measure the difference between local and non-local domestic banks. The results in Panel D confirm our previous findings in Panels A and B relative to local domestic banks and additionally indicate that differences in maturities between local domestic banks and foreign subsidiaries become more pronounced as geographical distance increases, confirming the importance of this disciplinary tool for foreign subsidiaries. As before, collateral seems to play a more important role for foreign branches (i.e., a non-local foreign branch is 44.4 percent more likely to have collateral relative to a domestic local bank, while a non-local foreign subsidiary is only 11.4 percent more likely). Interestingly, geographical distance seems to make no difference in the relative use of collateral and maturity for domestic banks.

1.4.5 Bank Ownership and Ex Post Loan Performance

In this final section, we compare the ex post performance of domestic and foreign bank loans. If foreign and domestic banks set loan conditions and price loans in an optimal way, we should not observe any systematic differences in the ex post performance of their loans, especially when holding differences in their clienteles fixed. We study three measures of ex post loan performance: *arrears or defaults* and *net return on loans* and present OLS regressions that compare the ex post performance of loans originated by domestic and foreign banks to the *same firm* in the *same month* for all loans in our sample, for secured and unsecured loans separately, for loans with maturities below and above the sample median, and for secured and

(2005).

[28] *Foreign Branch*is not interacted with $Non-Local$ as all loans originated by foreign branches are non-local as foreign branches have the headquarters are outside Bolivia. Domestic banks and foreign subsidiaries instead can be either local or non-local depending on whether the loan was originated in the region whether the bank is headquartered.

unsecured loans each with shorter and longer maturities.

First, we define a dummy variable, *Arrears* or *Default*, that equals one if a loan is in arrears for more than 30 days or if it is downgraded to the default status (rating 5), and equals zero otherwise. Regressions in Panel A of Table 1.11 show that loans originated by foreign banks to the same firm in the same month are 3.5 percentage points more likely to go in arrears or default, consistent with foreign banks facing higher enforcement costs. This finding is driven by unsecured and longer maturity loans.[29] In particular, we find no significant differences in repayment for secured and for short-term loans; significant differences are only found for unsecured and for long-term loans. When splitting the sample in two ways, the higher incidence of repayment problems for foreign bank loans is only confirmed for loans that are both unsecured and of longer maturity. The results thus suggest that foreign banks are able to effectively mitigate credit risk problems using collateral and shorter maturities. In unreported robustness checks, we also confirm these results with descriptive statistics or corresponding specifications with firm-fixed effects (instead of firm-time fixed effects) to alleviate possible concerns that insignificant differences in some sub-samples are due to exhaustive fixed effects saturating variation.

While foreign banks seem to experience higher arrears or defaults when they lend without collateral and at longer maturities, they still might be able to recover loans later on. Hence, we also compute the bank's net returns on each loan. Given the systematic differences in contract characteristics between domestic and foreign bank loans and differences in funding costs between domestic and foreign banks, the bank's net return on loans provides a more comprehensive measure of loan performance. Following Haselmann, Schoenherr and Vig (2013) and Skrastins and Vig (2014) we define the gross return on a loan (ROL) to firm i from bank j for the entire loan spell as:

$$\text{ROL}_{ij} = \sum_{t=1}^{T} \frac{\text{Loan Balance}_{ijt}}{\sum_{t=1}^{T} \text{Loan Balance}_{ijt}} \left[(1 - \mathbb{1}_{\{NPL=1\}}) r_{ijt} + \mathbb{1}_{\{NPL=1\}} \text{Loss}_{ijt} \right], \tag{1.4.1}$$

where the first term stands for the ratio of the outstanding loan amount to firm i from bank

[29] Note that results also hold when we include control variables as in previous regressions as well as the interest rate.

j at the beginning of period t ($LoanBalance_{ijt}$) to the sum of the outstanding loan amounts over the loan spell. The indicator function equals one when a loan has overdue payments between$t$ and $t+1$, r_{ijt} is the interest rate on the loan, and $Loss_{ijt}$ is the loss of the bank. We calculate the loss of the bank as the written-off amount or the overdue amount over the contract amount. The weights ensure that returns or defaults at the beginning of the loan spell receive more weight than those at the end of the loan spell when most of the loan has been repaid. To account for funding costs differences between domestic and foreign banks, we further subtract from the *ROL* each bank's cost of deposits at loan origination. We refer to this as the *Net Return on Loans* and use it as our main measure of returns on loans in Panel B.

The results in Panel B show that on average foreign banks have 1.26 percentage points lower net returns on their loans than domestic banks.[30] This is a more pronounced for unsecured loans (1.73 percentage points) as opposed to secured loans (insignificant) and for longer maturity (1.69 percentage points) than shorter maturity loans (0.82 percentage points). When splitting the sample, we find no significant difference in their net returns on loans that are both secured and of short maturity and the highest difference for loans that are both unsecured and of longer maturity (2.06 percentage points). In unreported robustness checks, we also confirm the results in Panel C using alternatively the gross return on loan (*ROL*) or a gross return on loans using discounted weights as in Haselmann et al. (2013).[31]

In summary, we find that foreign banks have a higher incidence of arrears or defaults and lower returns on their loans than domestic banks, consistent with foreign banks being at an informational and enforcement disadvantage relative to domestic banks, even when lending to the same firm at the same time. Collateral and maturity are found to help foreign banks overcome these problems.

1.4.6 Additional Robustness Tests

Next, we subject our results to additional robustness checks. (To conserve space, the results of some of these checks are discussed below, but are not presented in tables.) First, we investigate

[30]Notice that the number of observations reduces form 5,137 to 5,102 due to missing values for bank funding costs.

[31]Haselmann, et al. (2013) discount the weights in equation (1.4.1) to account for the time value of money.

whether within a given month there is any systematic pattern in the order in which firms obtain loans from domestic and foreign banks and whether our findings are sensitive to any such ordering. For the 5,137 loans in our sample, we find that in 2,330 cases the domestic bank was first, in 2,417 cases the foreign bank was first, and in 390 cases the foreign and the domestic banks granted the loans on exactly the same day. Re-estimating equation (1.3.1) for these three sub-samples yields results that are similar with those presented in Tables 1.3 and 1.4 and detects no statistically significant differences between them.

Second, we explore whether our results vary depending on whether multiple loans to the same firm within a given month were originated on the same or a different day. Loans originated on the same day could carry systematically different contract terms if such loans are part of some informal syndicated scheme. This could explain some of the observed differences in the loan contract terms between domestic and foreign bank loans if domestic and foreign banks were more or less likely to engage in such informal syndicates. Re-estimating an augmented version of equation (1.3.1) that allows for an interaction term between each explanatory variable and a dummy variable that indicates whether all loans to the same firm in a given month were originated at the same day, reveals that there are no statistically significant differences between them and our results are confirmed for both sets of loans.

In a final test, we examine whether the differences between domestic and foreign banks documented earlier are partly explained by the fact that the average size of foreign banks is much larger than the average size of domestic banks.[32] We therefore add the natural logarithm of the average total assets of banks—measured on a global level— as an additional explanatory variable to our specifications of equation (1.3.1). As can be observed in Table 1.12, in the collateral and interest rate regressions bank size produces similar results as foreign ownership, suggesting that larger size may indeed be part of the reason foreign banks behave differently. Relative to previous estimates, some of the differences between domestic and foreign banks are marginally absorbed by differences in bank size. Nevertheless, in all specifications, the

[32]Stein (2002) considers both bank size and foreign ownership as measures of hierarchical distance within an organization that are likely to generate differences in lending technologies. Empirically, Berger, et al. (2001) and Berger, et al. (2005) show that large banks, similar to foreign banks, target rather large and transparent firms. For a sample of Latin American countries Clarke, et al. (2005) find that large foreign banks lend more to small and medium sized companies than large domestic banks, while they find the reverse relationship for small banks.

coefficients of the foreign bank dummy retain their signs and statistical significance, suggesting that differences between domestic and foreign banks go beyond their mere differences in size.

One concern with these results is that there is almost no overlap in the distribution of bank size between foreign and domestic banks: only one foreign bank has a size that lies between the sizes of the smallest and largest domestic banks and three of the six foreign banks are more than 10 times larger than the largest domestic bank. Hence, in unreported regressions available on request, we re-estimate our specifications in Table 1.12 using a sub-sample of domestic and foreign banks in which the two groups are not significantly different in size. In particular, we exclude the four largest foreign banks (Citibank, ABN Amro, Banco do Brasil, and Banco de la Nacion Argentina) and confirm our previous findings.[33]

1.5 Conclusions

With the increase of the worldwide globalization of financial markets, the assessment of foreign bank entry and presence has become an important question for researchers and policymakers alike. The effects of foreign bank participation are not only important for the development of the banking sector in a country but will also have real effects on the economy as a whole, especially in countries with bank-finance dependent firms. Most country-level and cross-country studies confirm that foreign banks tend to lend to large and transparent firms and thus "cherry-pick" clients, leaving the difficult firms to domestic banks. Fewer studies find that foreign and large banks engage in lending to both smaller and larger firms.

Holding differences in their borrower clienteles constant, our study contributes to this literature by showing that foreign banks use contract design and credit scoring models to overcome distance constraints emanating from their size and steeper organizational structures and that such mechanisms are effective in terms of ex post loan performance. Such mechanisms however may carry important limitations. The stronger reliance on collateral and short maturity loans may reduce the options for firm investment. It may also have important repercussions from a capital allocation perspective, as it implies a shift towards firms with short-term financing needs

[33]Similar results are obtained if we instead exclude the three largest foreign banks (i.e., Citibank, ABN Amro, and Banco do Brasil) and the two smallest domestic banks (i.e., Banco Ganadero and Banco Economico).

and with pledgeable assets both on the intensive and extensive margin, limiting the potentially beneficial role foreign banks can play in developing countries.

References

Agarwal, Sumit, and Robert Hauswald, 2010, Distance and private information in lending, *Review of Financial Studies* 23, 2757-2788.

Aghion, P., and J. Tirole, 1997, Formal and real authority in organizations, *Journal of Political Economy* 105, 1-29.

Aghion, Philippe, and Patrick Bolton, 1997, A theory of trickle-down growth and development, *Review of Economic Studies* 64, 151-172.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2005, Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools, *Journal of Political Economy* 113, 151-184.

Ayyagari, Meghana, Thorsten Beck, and Asli Demirgüç-Kunt, 2007, Small and medium enterprises across the globe, *Small Business Economics* 29, 415-434.

Barnea, Amir, Robert A. Haugen, and Lemma W. Senbet, 1980, A rationale for debt maturity structure and call provisions in the agency theoretic framework, J*ournal of Finance* 35, 1223-1234.

Beck, Thorsten, Patrick Behr, and Andreas Madestam, 2014, Sex and credit: Is there a gender bias in lending?, European Banking Center Discussion Paper No. 2012-017.

Beck, Thorsten, Asli Demirgüç-Kunt, and María Soledad Martínez Pería, 2011, Bank financing for SMEs: Evidence across countries and bank ownership types, *Journal of Financial Services Research* 39, 35-54.

Beck, Thorsten, and María Soledad Martínez Pería, 2010, Foreign bank participation and outreach: Evidence from Mexico, *Journal of Financial Intermediation* 19, 52-73.

Bekaert, Geert, Campbell R. Harvey, and Christian T. Lundblad, 2005, Does financial liberalization spur growth?, *Journal of Financial Economics* 77, 3-55.

Bellows, John, and Edward Miguel, 2009, War and local collective action in Sierra Leone, *Journal of Public Economics* 93, 1144-1157.

Berger, Allen N., Scott W. Frame, and Vasso Ioannidou, 2011, Tests of ex ante versus ex post theories of collateral using private and public information, *Journal of Financial Economics* 100, 85-97.

Berger, Allen N., Leora F. Klapper, María Soledad Martínez Pería, and Rida Zaidi, 2008, Bank ownership type and banking relationships, *Journal of Financial Intermediation* 17, 37-62.

Berger, Allen N., Leora F. Klapper, and Gregory F. Udell, 2001, The ability of banks to lend to informationally opaque small businesses, *Journal of Banking & Finance* 25, 2127-2167.

Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein, 2005, Does function follow organizational form? Evidence from the lending practices of large and small banks, *Journal of Financial Economics* 76, 237-269.

Berger, Allen N., Richard J. Rosen, and Gregory F. Udell, 2007, Does market size structure affect competition? The case of small business lending, *Journal of Banking & Finance* 31,

11-33.

Berger, Allen N., and Gregory F. Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of Business* 68, 351-81.

Berger, Allen N., and Gregory F. Udell, 2006, A more complete conceptual framework for SME finance, *Journal of Banking & Finance* 30, 2945-2966.

Berlin, Mitchell, and Loretta J. Mester, 1992, Debt covenants and renegotiation, *Journal of Financial Intermediation* 2, 95-133.

Besanko, David, and Anjan V. Thakor, 1987, Collateral and rationing: Sorting equilibria in monopolistic and competitive credit markets, *International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association* 28, 671-89.

Bester, Helmut, 1985, Screening vs. rationing in credit markets with imperfect information, *The American Economic Review* 75, 850-855.

Bester, Helmut, 1987, The role of collateral in credit markets with imperfect information, *European Economic Journal* 31, 887-899.

Boot, A. W. A., A. V. Thakor, and G. F. Udell, 1991, Secured lending and default risk - equilibrium-analysis, policy implications and empirical results, *Economic Journal* 101, 458-472.

Boot, Arnoud W. A., and Anjan V. Thakor, 1994, Moral hazard and secured lending in an infinitely repeated credit market game, *International Economic Review* 35, 899-920.

Bruno, Valentina, and Robert Hauswald, 2014, The real effect of foreign banks, *Review of Finance* 18, 1683-1716.

Cameron, A. Colin , and Pravin K. Trivedi, 2005. *Microeconometrics: Methods and Applications* (Cambridge University Press, New York).

Cerqueiro, Geraldo, Hans Degryse, and Steven Ongena, 2011, Rules versus discretion in loan rate setting, *Journal of Financial Intermediation* 20, 503-529.

Cerutti, Eugenio, Giovanni Dell'Ariccia, and Maria Soledad Martínez Pería, 2007, How banks go abroad: Branches or subsidiaries?, *Journal of Banking & Finance* 31, 1669-1692.

Claessens, Stijn, and Neeltje van Horen, 2014, Foreign banks: Trends and impact, *Journal of Money Credit and Banking* 46, 295-326.

Clarke, George R. G., Robert Cull, María Soledad Martínez Pería, and Susana M. Sanchez, 2005, Bank lending to small businesses in Latin America: Does bank origin matter?, *Journal of Money, Credit and Banking* 37, 83-118.

De Haas, Ralph, Daniel Ferreira, and Anita Taci, 2010, What determines the composition of banks' loan portfolios? Evidence from transition countries, *Journal of Banking & Finance* 34, 388-398.

De la Torre, Augusto, María Soledad Martínez Pería, and Sergio L. Schmukler, 2010, Bank involvement with SMEs: Beyond relationship lending, *Journal of Money, Credit and Banking* 34, 2280-2293.

Degryse, Hans, Vasso Ioannidou, and Erik von Schedvin, 2015, On the Non-Exclusivity of Loan Contracts: An Empirical Investigation, *Management Science*, forthcoming.

Degryse, Hans, and Steven Ongena, 2005, Distance, lending relationships, and competition, *Journal of Finance* 60, 231-266.

Detragiache, Enrica, Poonam Gupta, and Thierry Tressel, 2008, Foreign banks in poor countries: Theory and evidence, *Journal of Finance* 63, 2123-2160.

Diamond, Douglas W., 1991, Debt maturity structure and liquidity risk, *Quarterly Journal of Economics* 106, 709-737.

Elsas, Ralf, and Jan Pieter Krahnen, 1998, Is relationship lending special? Evidence from credit-file data in Germany, *Journal of Banking & Finance* 22, 1283-1316.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig, 2012, Cultural proximity and loan outcomes NBER working papers, 18096 (NBER, Cambridge, Massachusetts, USA).

Flannery, Mark J., 1986, Asymmetric information and risky debt maturity choice, *Journal of Finance* 41, 19-37.

Flannery, Mark J., 1996, Financial crises, payment system problems, and discount window lending, *Journal of Money Credit and Banking* 28, 804-824.

Giannetti, Mariassunta, and Steven Ongena, 2009, Financial integration and firm performance: Evidence from foreign bank entry in Emerging Markets, *Review of Finance* 13, 181-223.

Giannetti, Mariassunta, and Steven Ongena, 2012, "Lending by example": Direct and indirect effects of foreign banks in emerging markets, *Journal of International Economics* 86, 167-180.

Gormley, Todd A., 2010, The impact of foreign bank entry in Emerging Markets: Evidence from India, *Journal of Financial Intermediation* 19, 26-51.

Harhoff, Dietmar, and Timm Körting, 1998, Lending relationships in Germany - Empirical evidence from survey data, *Journal of Banking & Finance* 22, 1317-1353.

Hart, Oliver, and John Moore, 1994, A theory of debt based on the inalienability of human capital, *Quarterly Journal of Economics* 109, 841-879.

Haselmann, Rainer, Katharina Pistor, and Vikrant Vig, 2010, How law affects lending, *Review of Financial Studies* 23, 549-580.

Haselmann, Rainer, David Schoenherr, and Vikrant Vig, 2013, Lending in social networks, Working Paper.

Henry, Peter B. , 2000, Do stock market liberalizations cause investment booms?, *Journal of Financial Economics* 58, 301-344.

Holmstrom, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, *Quarterly Journal of Economics* 112, 663-691.

Ioannidou, Vasso, and Steven Ongena, 2010, "Time for a change": Loan conditions and bank behavior when firms switch banks, *Journal of Finance* 65, 1847-1877.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? Evidence from subprime loans, *Quarterly Journal of Economics* 125,

307-362.

Liberti, Jose M, and Atif R Mian, 2009, Estimating the effect of hierarchies on information use, *Review of Financial Studies* 22, 4057-4090.

Machauer, Achim, and Martin Weber, 1998, Bank behavior based on internal credit ratings of borrowers, *Journal of Banking & Finance* 22, 1355-1383.

Mian, Atif, 2006, Distance constraints: The limits of foreign lending in poor economies, *Journal of Finance* 61, 1465-1505.

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147-175.

Ortiz-Molina, Hernan, and Maria F. Penas, 2008, Lending to small businesses: the role of loan maturity in addressing information problems, *Small Business Economics* 30, 361-383.

Petersen, Mitchell A., 2004, Information: hard and soft, (Kellog School of Management, Northwestern University, and NBER).

Petersen, Mitchell A., and Raghuram G. Rajan, 1994, The benefits of lending relationships: Evidence from Small Business Data, *Journal of Finance* 49, 3-37.

Qian, J., and P. E. Strahan, 2007, How laws and institutions shape financial contracts: The case of bank loans, *Journal of Finance* 62, 2803-2834.

Rajan, Raghuram G., 1992, Insiders and outsiders: The choice between informed and arm's-length debt, *Journal of Finance* 47, 1367-400.

Rajan, Uday, Amit Seru, and Vikrant Vig, 2015, The failure of models that predict failure: Distance, incentives, and defaults, *Journal of Financial Economics* 115, 237-260.

Roberts, Michael R. , and Toni M. Whited, 2012, Endogeneity in empirical corporate finance, in Milton Harris and Rene M. Stulz George M. Constantinides, ed.: *Handbook of the Economics of Finance* (Elsevier).

Sirtaine, Sophie, Ilias Skamnelos, and Sissy Frank, 2004, Bolivia: Challenges in the corporate and banking sectors, Working Paper (World Bank, Washington DC).

Skrastins, Janis, and Vikrant Vig, 2014, How organizational hierarchy affects information production, Working Paper.

Stein, Jeremy C. , 2002, Information production and capital allocation: Decentralized versus hierarchical firms, *Journal of Finance* 57, 1891-1921.

Tables

Table 1.1: Summary Statistics for Commercial Banks Operating in Bolivia

The Table provides summary statistics for all commercial banks that were active in Bolivia between March 1999 and December 2003. We distinguish between foreign branches, foreign subsidiaries and domestic banks. A bank is considered to be foreign if more that 50% of its shares are foreign-owned**. Entry/ Acquisition** indicates at which point in time the bank entered the market or was acquired by a foreign bank. **Total Assets** stands for the average value of total assets in millions of US$ during the sample period. **Market Share** stands for a bank's total loans in the country to the total loans in the country per month. **Cost of Deposits** stands for the average interest rate on dollar deposits in a month.

Bank	**Entry/**	**Total Assets**		**Market Share**		**Cost of Deposits**	
	Acquisition	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
Foreign Branches							
Citibank*	01/01/66	975134.000	198429.000	0.069	0.024	3.034	1.634
ABN Amro*	07/08/98	530089.000	48568.000	0.014	0.007	4.721	0.454
Banco do Brasil*	07/01/61	76124.000	16741.000	0.018	0.012	3.938	1.324
Banco de la Nación Argentina*	04/28/58	14614.171	4031.653	0.016	0.012	5.320	1.361
Foreign Subsidiaries							
Banco Santa Cruz	07/17/98	833.479	354.155	0.105	0.044	3.003	1.686
Banco de Crédito de Bolivia	12/30/92	589.057	97.402	0.161	0.053	4.245	1.438
Banco Solidario	03/15/99	94.936	6.970	0.004	0.002	5.509	1.850
Domestic							
Banco Industrial		682.490	48.450	0.263	0.054	4.021	1.366
Banco Nacional de Bolivia		621.065	17.808	0.118	0.047	5.037	1.434
Banco Mercantil		598.895	31.616	0.108	0.038	4.691	1.514
Banco de la Unión		443.784	90.026	0.060	0.027	5.886	1.766
Banco Económico		284.716	36.613	0.039	0.020	6.265	1.503
Banco Ganadero		207.390	21.950	0.046	0.016	5.586	1.599

* consolidated total assets based on annual reports from December 1998 until December 2003.

Table 1.2: Summary Statistics

The Table reports summary statistics for the entire sample of 32,279 loans and 2,672 firms, referred to as "Universe", during the period March 1999 to December 2003. Summary statistics are also provided separately for loans originated by foreign and domestic banks. The definitions of all variables can be found in Table A.1 in the Appendix. The stars next to the mean values of domestic bank loans under "Universe" indicate whether the differences between domestic and foreign banks are statistically significant using the *t*-test. The second part of the table "Sample", reports summary statistics for a sub-sample of 5,137 loans to 287 firms and compares it to the "Universe" of all loans. Like in the first part, statistics are provided for all loans in the sample as well as for foreign and domestic banks separately. The stars next to each mean value indicate whether it is statistically different from its corresponding value for the entire sample using a sample adjusted t-test. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variable Names	Universe						Sample					
	All		Foreign		Domestic		All		Foreign		Domestic	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Loan Terms												
Installment	0.470	0.499	0.500	0.500	0.450 ***	0.498	0.461	0.499	0.482	0.500	0.437	0.496
Amount	161,908	468,898	204,725	564,162	134,714 ***	394,297	251,098 ***	524,793	239,967 ***	436,543	263,231 ***	606,387
Collateral	0.245	0.430	0.376	0.484	0.161 ***	0.368	0.331 ***	0.470	0.461 ***	0.499	0.188 ***	0.391
Maturity	10.859	16.272	8.304	10.926	12.481 ***	18.717	8.817 ***	11.211	7.049 ***	7.382	10.744 ***	14.007
Interest Rate	13.448	2.887	13.041	3.020	13.706 ***	2.769	12.617 ***	2.951	12.203 ***	2.863	13.069 **	2.979
Loan Spread	9.949	2.763	9.352	2.990	10.328 ***	2.537	9.155 ***	2.594	8.774 ***	2.654	9.570	2.460
Legal Structure												
Sole Proprietorship	0.125	0.331	0.096	0.295	0.144 ***	0.351	0.046 ***	0.210	0.034 ***	0.180	0.060 ***	0.237
General Partnership	0.009	0.096	0.005	0.073	0.012 ***	0.108	0.005 ***	0.070	0.001 ***	0.039	0.009 *	0.092
Limited Partnership	0.130	0.337	0.139	0.346	0.125 ***	0.331	0.147 ***	0.354	0.166 ***	0.373	0.125	0.331
Joint Stock Company	0.229	0.420	0.273	0.446	0.201 ***	0.401	0.358 ***	0.479	0.381 ***	0.486	0.332 ***	0.471
Limited Liability Company	0.486	0.500	0.472	0.499	0.494 ***	0.500	0.428 ***	0.495	0.411 ***	0.492	0.446 ***	0.497
Other	0.020	0.142	0.014	0.118	0.025 ***	0.155	0.017 **	0.128	0.006 ***	0.077	0.028	0.166
Bank Debt												
Outstanding Debt	1,991,796	3,879,224	2,410,193	4,194,117	1,726,061 ***	3,640,433	5,452,792 ***	6,474,100	5,146,245 ***	6,395,487	5,786,901 ***	6,543,670
Credit Quality												
Past Non-Performance	0.209	0.407	0.246	0.431	0.186 ***	0.389	0.304 ***	0.460	0.284 ***	0.451	0.325 ***	0.468
Rating 1	0.873	0.332	0.860	0.347	0.882 ***	0.323	0.857	0.350	0.815 ***	0.388	0.903	0.296
Rating 2	0.098	0.298	0.096	0.295	0.100	0.299	0.119 ***	0.324	0.150 ***	0.357	0.085 ***	0.279
Rating 3	0.024	0.154	0.035	0.185	0.017 ***	0.130	0.023	0.150	0.033	0.179	0.012 **	0.110
Rating 4	0.004	0.063	0.008	0.091	0.001 ***	0.036	0.001 ***	0.028	0.001 ***	0.039	0 ***	0
Relationship Characteristics												
Multiple Relationships	0.555	0.497	0.620	0.485	0.514 ***	0.500	1	0	1	0	1	0
Rel Duration	22.079	16.065	20.840	15.272	22.866 ***	16.500	23.071 ***	16.354	23.407 ***	16.543	22.705	16.142
Scope	0.259	0.438	0.224	0.417	0.281 ***	0.450	0.255	0.436	0.206 **	0.404	0.310 ***	0.462
Primary	0.714	0.452	0.665	0.472	0.744 ***	0.436	0.283 ***	0.451	0.312 ***	0.464	0.252 ***	0.434
Observations	32,279		12,538		19,741		5,137		2,679		2,458	

Table 1.3: Determinants of Collateral and Maturity

The Table reports OLS regressions for a sample of 5,137 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. The dependent variables are *Collateral*, a dummy variable indicating whether the loan contract includes the pledge of collateral and *Maturity*, the natural logarithm of one plus the number of months between loan origination and maturity. Columns I-V (VI-X) report regression results for *Collateral* (*Maturity*) as the dependent variable, with different control variables and for smaller and larger firms separately. Columns IV-V and IX-X report regression results for collateral and maturity for sub-samples of firms with outstanding bank debt below or above the sample median, denoted as *Smaller Firms* and *Larger Firms*, respectively. Definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (presented in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III	IV	V	VI	VII	VIII	IX	X
	Collateral					Maturity				
	All firms			Smaller Firms	Larger Firms	All firms			Smaller Firms	Larger Firms
Bank Characteristics										
Foreign Bank	0.265***	0.272***	0.277***	0.325***	0.263***	-0.188***	-0.191***	-0.266***	-0.252***	-0.269***
	(0.041)	(0.038)	(0.039)	(0.055)	(0.047)	(0.046)	(0.043)	(0.042)	(0.075)	(0.049)
Firm Characteristics										
Rating 2		0.154***	0.139**	0.185	0.126**		-0.016	-0.023	-0.046	-0.026
		(0.059)	(0.059)	(0.122)	(0.063)		(0.134)	(0.094)	(0.168)	(0.108)
Ratings 3 & 4		-0.167*	-0.182*	-0.081	-0.225**		0.008	-0.207	0.081	-0.318
		(0.090)	(0.097)	(0.193)	(0.114)		(0.246)	(0.186)	(0.391)	(0.194)
Relationship Characteristics										
Rel Duration		-0.094***	-0.094***	-0.036	-0.122***		-0.231**	-0.174	-0.130	-0.186
		(0.027)	(0.026)	(0.044)	(0.031)		(0.113)	(0.121)	(0.184)	(0.148)
Rel Duration-Square							0.047*	0.037	0.036	0.036
							(0.027)	(0.032)	(0.049)	(0.038)
Rel Scope		-0.059	-0.058	0.073	-0.083*		-0.015	-0.117**	-0.159	-0.105*
		(0.044)	(0.044)	(0.076)	(0.049)		(0.066)	(0.054)	(0.110)	(0.061)
Primary Bank		-0.118***	-0.123***	-0.132**	-0.121***		0.222***	0.121***	0.187***	0.082
		(0.036)	(0.034)	(0.058)	(0.043)		(0.059)	(0.043)	(0.071)	(0.053)
Other Contract Terms										
Installment			-0.049	-0.057	-0.050			0.846***	0.977***	0.812***
			(0.041)	(0.055)	(0.047)			(0.075)	(0.109)	(0.088)
Loan Amount			0.028**	0.064**	0.023			0.063**	0.103**	0.057**
			(0.013)	(0.025)	(0.014)			(0.025)	(0.048)	(0.028)
Constant	0.192***	0.493***	0.200	-0.365	0.354*	2.059***	2.243***	1.155***	0.499	1.303***
	(0.021)	(0.078)	(0.165)	(0.280)	(0.182)	(0.024)	(0.118)	(0.314)	(0.589)	(0.355)
Fixed Effects										
Firm×Time Fixed Effects	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
R-squared	0.508	0.537	0.540	0.640	0.519	0.447	0.459	0.632	0.701	0.613
Observations	5,137	5,137	5,137	1,129	4,008	5,137	5,137	5,137	1,129	4,008

Table 1.4: Determinants of Loan Interest Rate

The Table reports OLS regressions for a sample of 5,137 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. The dependent variable is *Loan Spread*, the loan interest rate minus the rate on US Treasury securities of comparable maturity at loan origination. Columns I-III report regression results with different control variables. Columns IV-V report regression results with all control variables, where the sample is divided between firms with outstanding bank debt below or above the sample median, denoted as *Smaller Firms* and *Larger Firms*, respectively. Columns VI-X report augmented specifications of those presented in Columns I-V where bank funding costs and market shares are added among the control variables. The definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (presented in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III	IV	V	VI	VII	VIII	IX	X
	Benchmark Specifications					Funding Costs & Market Share				
	All firms			Smaller Firms	Larger Firms	All firms			Smaller Firms	Larger Firms
Bank Characteristics										
Foreign Bank	-0.887***	-0.893***	-1.068***	-0.961***	-1.126***	-0.689***	-0.729***	-0.987***	-0.692**	-1.101***
	(0.134)	(0.135)	(0.147)	(0.280)	(0.168)	(0.179)	(0.173)	(0.171)	(0.343)	(0.193)
Cost of Deposits (%)						0.107	0.082	0.064	0.185	0.045
						(0.076)	(0.072)	(0.060)	(0.136)	(0.067)
Market Share						1.499	1.556	-0.201	2.309	-1.012
						(0.954)	(0.965)	(0.986)	(1.742)	(1.133)
Firm Characteristics										
Rating 2		0.468*	0.513**	0.067	0.632**		0.500*	0.511**	0.138	0.623**
		(0.268)	(0.241)	(0.485)	(0.256)		(0.264)	(0.244)	(0.487)	(0.263)
Ratings 3 & 4		0.831	0.661	-0.429	1.099**		0.824	0.668	-0.411	1.118*
		(0.650)	(0.544)	(1.284)	(0.551)		(0.638)	(0.556)	(1.291)	(0.568)
Relationship Characteristics										
Rel Duration		0.393	0.119	1.270**	-0.337		0.365	0.112	1.253**	-0.351
		(0.291)	(0.243)	(0.596)	(0.255)		(0.291)	(0.246)	(0.583)	(0.256)
Rel Duration-Square		-0.096	-0.046	-0.307*	0.049		-0.089	-0.041	-0.299*	0.058
		(0.071)	(0.061)	(0.157)	(0.063)		(0.071)	(0.062)	(0.154)	(0.063)
Rel Scope		0.223	0.121	0.209	0.103		0.139	0.138	0.140	0.163
		(0.194)	(0.163)	(0.342)	(0.178)		(0.191)	(0.158)	(0.335)	(0.176)
Primary Bank		-0.491***	-0.326**	-0.614**	-0.147		-0.484***	-0.309**	-0.579**	-0.127
		(0.153)	(0.137)	(0.259)	(0.162)		(0.153)	(0.137)	(0.242)	(0.164)
Other Contract Terms										
Installment			0.542***	0.347	0.569***			0.519***	0.369	0.547***
			(0.149)	(0.376)	(0.159)			(0.151)	(0.367)	(0.162)
Loan Amount			0.017	0.094	0.003			0.021	0.078	0.012
			(0.098)	(0.115)	(0.108)			(0.101)	(0.116)	(0.111)
Collateral			-0.371**	-0.133	-0.428**			-0.393**	-0.110	-0.467**
			(0.162)	(0.265)	(0.181)			(0.161)	(0.276)	(0.184)
Maturity			-1.191***	-1.213***	-1.164***			-1.188***	-1.172***	-1.170***
			(0.133)	(0.187)	(0.160)			(0.134)	(0.190)	(0.160)
Constant	9.617***	9.365***	11.783***	11.299***	12.010***	8.852***	8.749***	11.432***	10.070***	11.833***
	(0.070)	(0.280)	(0.954)	(1.397)	(1.038)	(0.445)	(0.468)	(1.046)	(1.779)	(1.140)
Fixed Effects										
Firm×Time Fixed Effects	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
R-squared	0.671	0.678	0.731	0.734	0.718	0.673	0.680	0.731	0.737	0.719
Observations	5,137	5,137	5,137	1,129	4,008	5,131	5,131	5,131	1,126	4,005

Table 1.5: Contract Terms for Opposite Sample

The Table reports OLS regressions for a sample of 27,142 loans to 2,671 firms that received a new loan from foreign or from domestic banks, but not from both, in the same month during the period March 1999 to December 2003. Columns I-III report specifications of the *Collateral*, *Maturity*, and *Loan Spread* models with all control variables. *Collateral* is a dummy variable indicating whether the loan contract includes the pledge of collateral. *Maturity* is the natural logarithm of one plus the number of months between loan origination and maturity and *Loan Spread* is the loan interest rate minus the rate on US Treasury securities of comparable maturity at loan origination. The definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (presented in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III
	Collateral	Maturity	Loan Spread
Bank Characteristics			
Foreign Bank	0.185***	-0.338***	-1.045***
	(0.027)	(0.034)	(0.112)
Firm Characteristics			
Rating 2	-0.002	-0.061**	0.350***
	(0.021)	(0.031)	(0.108)
Ratings 3 & 4	0.069*	-0.054	0.719***
	(0.037)	(0.077)	(0.219)
Relationship Characteristics			
Rel Duration	-0.032***	-0.058**	-0.260***
	(0.008)	(0.028)	(0.091)
Rel Duration-Square		0.005	0.086***
		(0.009)	(0.030)
Rel Scope	-0.020*	-0.032*	0.041
	(0.012)	(0.017)	(0.065)
Primary Bank	-0.039**	0.069***	-0.408***
	(0.015)	(0.023)	(0.092)
Other Contract Terms			
Installment	0.053***	1.031***	0.641***
	(0.012)	(0.026)	(0.080)
Loan Amount	0.037***	0.123***	0.003
	(0.005)	(0.012)	(0.044)
Collateral			-0.254***
			(0.075)
Maturity			-1.194***
			(0.050)
Constant	-0.243***	0.542***	12.829***
	(0.063)	(0.129)	(0.436)
Fixed Effects			
Firm Fixed Effects	Included	Included	Included
Time Fixed Effects	Included	Included	Included
R-squared	0.507	0.688	0.718
Observations	27,142	27,142	27,142

Table 1.6: Matching of Contract Terms

The Table reports results from a matching exercise using our sample of 5,137 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. For each specification, we estimate the average differences between the contract terms of domestic and foreign bank loans using the sub-sample of loans that satisfy the matching criteria in each case. In Column I, we match on firm identity and the month of loan origination. In Column II, we additionally match on ratings and relationship characteristics. In Column III, we also match on other contract terms as in Tables 3 and 4. For each matched sample, we report results for collateral, maturity, and loan spread. For the collateral regression, the dependent variable equals the difference in *Collateral* between the matched foreign and domestic bank loans. The dependent variables for the maturity and loan spread regressions are calculated similarly using *Maturity* and *Loan Spread* (see Table A.1 in the Appendix for detailed definitions of *Collateral*, *Maturity*, and *Loan Spread* as well all other matching variables). For each dependent variable and sample, we estimate an OLS regression with only a constant term, clustering standard errors at the firm level. For each specification, we report the estimated constant term, the standard error, and the number of observations. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III
Collateral Difference	0.303 ***	0.322 ***	0.424 ***
	(0.008)	(0.020)	(0.030)
	5,077	798	368
Maturity Difference	-0.196 ***	-0.177 ***	-0.122 ***
	(0.015)	(0.030)	(0.032)
	5,077	798	368
Loan Spread Difference	-0.834 ***	-1.230 ***	-1.286 ***
	(0.039)	(0.030)	(0.153)
	5,077	798	252

Table 1.7: Changes in Contract Terms after a Foreign Bank Takeover

The Table reports the changes in the contract terms of a domestic bank, Banco Boliviano Americano, after it has been taken over by a foreign bank, Banco Credito de Bolivia, in May 1999 using a difference-in-difference analysis. To determine how contract terms change once the bank is taken over, we use a constant sample of firms that borrowed from the acquired bank both before and after the takeover. In particular, we identify firms that between January 1998 and the month prior to the merger received new loans from their initial bank (Banco Boliviano Americano). We match these loans with any loans that the same firm received from the acquiring bank in the two years after the acquisition, dropping firms that prior to the acquisition were already customers of the acquiring bank (Banco Credit de Bolivia). This defines our treatment group resulting in 85 loans and 140 pairs. The column "Treatment" reports the average differences in the contract terms of loans to the same firm (After – Before). The column "Control" reports the average differences in the contract terms of domestic (not taken over) banks to their customers over the same period as the treatment group. Like for the treatment group we use a constant set of firms and report within bank-firm changes in contract terms over the comparison period. The last column "Treatment – Control" compares the changes of the treatment group relative to the changes of the control group. The stars next to each average change indicate whether the reported change is statistically different from zero. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	After - Before		
	Treatment	Control	Treatment - Control = 0
Collateral Difference	0.157 **	-0.031 ***	0.188 ***
Maturity Difference	-2.112 *	3.301 ***	-5.413 ***
Loan Spread Difference	0.072	0.432 ***	-0.360 **

Table 1.8: Lending Technologies and Loan Pricing

The Table reports OLS regressions for a sample of 5,137 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. The dependent variable is *Loan Spread*, the loan interest rate minus the rate on US Treasury securities of comparable maturity at loan origination. Columns I-III report regression results with all control variables, where each variable is interacted with the *Foreign Bank* dummy. The Column *Domestic* reports the domestic bank coefficients, × *Foreign* reports the interaction coefficients, and *Cumulative* reports the foreign bank coefficients. Columns IV-VI and VII-IX report the same regression results, where the sample is divided between firms with outstanding bank debt below or above the sample median, denoted as *Smaller Firms* and *Larger Firms*, respectively. The definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (presented in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III	IV	V	VI	VII	VIII	IX
		All Firms			Smaller Firms			Larger Firms	
	Domestic	× Foreign	Cumulative	Domestic	× Foreign	Cumulative	Domestic	× Foreign	Cumulative
Firm Characteristics									
Rating 2	0.168	0.478	0.646***	0.334	-0.183	0.150	0.164	0.613	0.776***
	(0.480)	(0.534)	(0.240)	(0.913)	(1.095)	(0.599)	(0.483)	(0.550)	(0.270)
Ratings 3 & 4	-0.900	2.097**	1.197**	-2.518	3.440	0.922	-0.401	1.883*	1.482**
	(0.896)	(1.052)	(0.583)	(1.875)	(2.190)	(1.200)	(0.870)	(1.103)	(0.638)
Relationship Characteristics									
Rel Duration	0.299	-0.146	0.152	1.570**	-0.618	0.952	-0.323	0.172	-0.151
	(0.350)	(0.516)	(0.380)	(0.717)	(0.789)	(0.745)	(0.369)	(0.572)	(0.428)
Rel Duration-Square	-0.091	0.052	-0.039	-0.381**	0.127	-0.254	0.042	-0.008	0.035
	(0.093)	(0.113)	(0.081)	(0.187)	(0.175)	(0.179)	(0.099)	(0.125)	(0.091)
Rel Scope	-0.249	0.800***	0.551**	-0.155	1.107	0.952*	-0.261	0.763**	0.502*
	(0.192)	(0.277)	(0.238)	(0.495)	(0.681)	(0.498)	(0.199)	(0.299)	(0.260)
Primary Bank	-0.758***	0.840**	0.082	-0.895**	0.631	-0.264	-0.548**	0.753*	0.205
	(0.229)	(0.374)	(0.234)	(0.396)	(0.749)	(0.484)	(0.269)	(0.405)	(0.262)
Other Contract Terms									
Installment	0.610**	-0.150	0.460**	0.206	0.176	0.382	0.698**	-0.263	0.435*
	(0.281)	(0.390)	(0.210)	(0.507)	(0.654)	(0.447)	(0.320)	(0.436)	(0.221)
Loan Amount	-0.046	0.112	0.066	0.175	-0.109	0.066	-0.129	0.236**	0.106
	(0.126)	(0.094)	(0.088)	(0.124)	(0.211)	(0.189)	(0.146)	(0.114)	(0.090)
Collateral	-0.127	-0.302	-0.429**	0.076	-0.310	-0.234	-0.121	-0.347	-0.468**
	(0.257)	(0.296)	(0.172)	(0.500)	(0.749)	(0.407)	(0.286)	(0.326)	(0.191)
Maturity	-1.170***	-0.018	-1.188***	-1.254***	0.132	-1.122***	-1.134***	-0.074	-1.209***
	(0.194)	(0.251)	(0.158)	(0.287)	(0.437)	(0.284)	(0.222)	(0.292)	(0.189)
Constant	12.531***	-2.713**	9.818***	10.584***	0.065	10.649***	13.657***	-4.452***	9.205***
	(1.195)	(1.159)	(1.023)	(1.416)	(2.390)	(2.286)	(1.419)	(1.351)	(1.025)
Fixed Effects									
Firm×Time Fixed Effects		Included			Included			Included	
R-squared		0.738			0.744			0.727	
Observations		5,137			1,129			4,008	

Table 1.9: Hard Information Determinants of Loan Interest Rates

The Table reports OLS regressions for a sample of 5,137 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. In all specifications, the dependent variable is the *Loan Spread*, the loan interest rate minus the rate on US Treasury securities of comparable maturity at loan origination that is regressed on credit ratings and past performance measures. Columns I-II report regression results for foreign and domestic banks separately, while Columns III-VI additionally split these two specifications between firms with outstanding bank debt below or above the sample median for foreign and domestic banks, denoted as *Smaller Firms* and *Larger Firms*, respectively. The definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (presented in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III	IV	V	VI
			Smaller Firms		Larger Firms	
	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic
Rating 2	1.563***	0.633**	1.052***	0.589	1.769***	0.760**
	(0.321)	(0.291)	(0.368)	(0.535)	(0.362)	(0.313)
Ratings 3 & 4	3.579***	1.038	1.220**	-1.999**	4.123***	1.934***
	(0.872)	(0.709)	(0.613)	(1.006)	(1.012)	(0.705)
Collateral	-0.514*	-0.757***	0.310	-0.422	-0.732**	-0.727***
	(0.290)	(0.254)	(0.343)	(0.592)	(0.335)	(0.276)
Constant	8.652***	9.646***	9.580***	10.656***	8.379***	9.313***
	(0.186)	(0.171)	(0.266)	(0.226)	(0.224)	(0.186)
R-squared	0.104	0.019	0.030	0.017	0.155	0.027
Observations	2,679	2,458	565	564	2,114	1,894

Table 1.10: What Drives "Distance Constraints"?

The Table reports modified specifications of the *Collateral*, *Maturity*, and *Loan Spreads* models. Panel A distinguishes foreign banks into foreign branches and foreign subsidiaries using specifications equivalent to those reported in Columns III and VII of Table 3 and Column III of Table 4 with firm×time fixed effects. The model is estimated with OLS using a sample of 689 loans to 30 firms that received a new loan from at least one domestic bank, one foreign branch and one foreign subsidiary in the same month during the period March 1999 to December 2003 (Restrictive Sample). Panel B relaxes the constraint that firms must be borrowing from all three types of banks in the same month and considers a sample of firms borrowing from the three types of banks at any point during the sample period, resulting in a sample of 7,040 loans to 117 firms (Larger Sample). The firm×time fixed effects are thus replaced with firm and time fixed effects. Panel C (D) a presents specifications similar to Panel B allowing for interactions with *Time Since Entry/Foreign Owned (Non-Local)*. The definitions of all variable are in Table A.1 in the Appendix. Standard errors, in parentheses, are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III
	Collateral	Maturity	Loan Spread
A. Branches vs. Subsidiaries - Restricted Sample			
Foreign Branch	0.422***	0.103	-1.633***
	(0.080)	(0.101)	(0.392)
Foreign Subsidiary	0.150*	-0.226***	-1.371***
	(0.086)	(0.077)	(0.297)
R-squared	0.466	0.509	0.706
Observations	689	689	689
T-test Branches vs. Subsidiaries (p-values)	0.018	0.022	0.457
B. Branches vs. Subsidiaries - Larger Sample			
Foreign Branch	0.438***	-0.209***	-1.819***
	(0.052)	(0.044)	(0.218)
Foreign Subsidiary	0.126***	-0.443***	-0.989***
	(0.040)	(0.059)	(0.181)
R-squared	0.326	0.507	0.603
Observations	7,040	7,040	7,040
T-test Branches vs. Subsidiaries (p-values)	0.000	0.000	0.000
C. Time since Entry/Foreign Owned - Larger Sample			
Foreign Branch	0.768***	-0.568***	-0.580
	(0.106)	(0.107)	(0.507)
Foreign Subsidiary	-0.147***	0.035	-0.277
	(0.052)	(0.064)	(0.297)
Foreign Branch×Time Since Entry	-0.100***	0.122***	-0.349**
	(0.032)	(0.031)	(0.143)
Foreign Subsidiary×Time Since Foreign Owned	0.167***	-0.286***	-0.440***
	(0.035)	(0.041)	(0.145)
R-squared	0.345	0.525	0.607
Observations	7,040	7,040	7,040
D. Geographical Distance - Larger Sample			
Foreign Branch	0.444***	-0.159***	-1.503***
	(0.054)	(0.046)	(0.263)
Foreign Subsidiary	0.143**	-0.305***	-1.019***
	(0.065)	(0.076)	(0.244)
Foreign Subsidiary×Non-Local	-0.028	-0.232**	0.143
	(0.081)	(0.089)	(0.337)
Non-Local	-0.013	-0.039	-0.590**
	(0.052)	(0.055)	(0.279)
R-squared	0.327	0.511	0.606
Observations	7,040	7,040	7,040
Foreign Subsidiary+Foreign Subsidiary×Non-Local	0.114**	-0.537***	-0.876***
	(0.049)	(0.071)	(0.253)

Table 1.11: Lending Techniques and Ex Post Loan Performance

The Table reports OLS regressions on the ex post loan performance for a sample of 5,137 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. In each panel, we report specifications using different measures of ex post loan performance as a dependent variable. In Panel A, we use a dummy variable *Arrears* or *Default* that equals one if a loan is in arrears for more than 30 days or is downgraded to the default status (rating 5) and zero otherwise. In Panel B, we use the net return on loans as indicated by equation (1.4.1) minus each bank's cost of deposit at loan origination. For each specification, we report the foreign bank coefficient, the constant term, the R-squared, and the number of observations. We estimate each model for all loans, for secured and unsecured loans separately, and for loans with maturities below and above the sample median. For *Ar*rears or *Defa*ults and *Net Returns on Loans*, we also estimate the model separately for secured and unsecured loans each with shorter and longer maturities including in all cases firm×time fixed effects. The definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (reported in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	All Loans	Secured	Unsecured	Shorter Maturity	Longer Maturity	Secured		Unsecured	
						Shorter Maturity	Longer Maturity	Shorter Maturity	Longer Maturity
A. Arrears or Default									
Foreign Bank	0.035 **	0.016	0.044 *	0.021	0.078 ***	0.030	0.016	0.034	0.077 *
	(0.015)	(0.025)	(0.025)	(0.015)	(0.026)	(0.030)	(0.027)	(0.029)	(0.042)
Constant	0.048 ***	0.068 ***	0.041 ***	0.041 ***	0.044 ***	0.039	0.093 ***	0.032 **	0.041 **
	(0.008)	(0.019)	(0.010)	(0.009)	(0.012)	(0.024)	(0.017)	(0.013)	(0.017)
Firm×Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.554	0.781	0.632	0.665	0.707	0.799	0.887	0.736	0.772
Observations	5,137	1,698	3,439	2,705	2,432	1,005	693	1,700	1,739
B. Net Return on Loans									
Foreign Bank	-1.260 ***	-0.687	-1.731 ***	-0.817 ***	-1.685 ***	0.117	-1.168 *	-1.396 ***	-2.056 ***
	(0.215)	(0.500)	(0.316)	(0.252)	(0.328)	(0.837)	(0.627)	(0.416)	(0.494)
Constant	11.197 ***	10.625 ***	11.470 ***	10.926 ***	11.416 ***	9.782 ***	11.221 ***	11.415 ***	11.512 ***
	(0.112)	(0.364)	(0.132)	(0.143)	(0.153)	(0.666)	(0.394)	(0.181)	(0.198)
Firm×Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.650	0.851	0.710	0.804	0.732	0.919	0.904	0.832	0.789
Observations	5,102	1,693	3,409	2,686	2,416	1,000	693	1,686	1,723

Table 1.12: Lending Technologies and Bank Size

The Table reports OLS regressions for a sample of 5,134 loans to 287 firms that received a new loan from at least one foreign and one domestic bank in the same month during the period March 1999 to December 2003. The sample size is smaller than in the benchmark regressions as bank size is not observed for all banks. *Bank Size* equals the natural logarithm of the originating bank's total assets. Columns I-III report augmented specifications of the *Collateral*, *Maturity,* and *Loan Spreads* models with all control variables where we add *Bank Size* among the bank characteristics. *Collateral* is a dummy variable indicating whether the loan contract includes the pledge of collateral. *Maturity* is the natural logarithm of one plus the number of months between loan origination and maturity and *Loan Spread* is the loan interest rate minus the rate on US Treasury securities of comparable maturity at loan origination. The definitions of all variables can be found in Table A.1 in the Appendix. Standard errors (presented in parentheses) are clustered at firm level. ***, **, and * indicate significance at the 1%, 5%, and 10%, respectively.

	I	II	III
	Collateral	Maturity	Loan Spread
Bank Characteristics			
Foreign Bank	0.195***	-0.376***	-0.743***
	(0.042)	(0.053)	(0.166)
Bank Size	0.032**	0.043***	-0.129***
	(0.013)	(0.010)	(0.041)
Firm Characteristics			
Rating 2	0.157***	-0.001	0.434**
	(0.057)	(0.097)	(0.217)
Ratings 3 & 4	-0.145	-0.158	0.541
	(0.096)	(0.175)	(0.519)
Relationship Characteristics			
Rel Duration	-0.099***	-0.132	0.009
	(0.026)	(0.119)	(0.251)
Rel Duration-Square		0.025	-0.011
		(0.031)	(0.065)
Rel Scope	-0.034	-0.085	0.036
	(0.040)	(0.054)	(0.158)
Primary Bank	-0.125***	0.119***	-0.316**
	(0.034)	(0.044)	(0.131)
Other Contract Terms			
Installment	-0.015	0.891***	0.360**
	(0.035)	(0.070)	(0.156)
Loan Amount	0.030**	0.067***	-0.000
	(0.012)	(0.024)	(0.100)
Collateral			-0.282*
			(0.150)
Maturity			-1.134***
			(0.128)
Constant	-0.039	0.798**	12.776***
	(0.173)	(0.320)	(0.995)
Fixed Effects			
Firm×Time Fixed Effects	Included	Included	Included
R-squared	0.551	0.640	0.737
Observations	5,134	5,134	5,134

Appendix

Table A.1: Definitions of Variables

Variable Names	Definitions
Loan Characteristics	
Installment	= 1 if an installment loan, and = 0 if a single payment loan.
Loan Amount	loan amount at loan origination in US dollars.
Collateral	= 1 if collateral was pledged at loan origination, and = 0 otherwise.
Maturity	number of months between loan origination and maturity.
Interest Rate	annual contractual interest rate at loan origination.
Loan Spread	loan interest rate minus rate on Treasury Bills of comparable maturity.
Geographical Location	= 1 if a loan has been issued in one of the 9 different regions of Bolivia or abroad, and = 0 otherwise: Chuquisaca, La Paz, Cochabamba, Oruro, Potosi, Tarija, Santa Cruz, Beni, Pando, USA, Argentina, Paraguay, Panama.
Legal Structure & Industry	
Sole Proprietorship	= 1 if the firms is a sole proprietorship, and = 0 otherwise.
General Partnership	= 1 if the firm is a general partnership (i.e., all partners have unlimited liability and ownership is not transferable), and = 0 otherwise.
Limited Partnership	= 1 if the firm is a limited partnership (i.e., some partners have limited liability and their ownership rights are transferable), and = 0 otherwise.
Joint Stock Company	= 1 if the firm is a joint-stock company (i.e., all partners have unlimited liability and their ownership rights are transferable) and = 0 otherwise.
Limited Liability Company	= 1 if the firm is a limited liability company (i.e., all partners have limited liability and transferable ownership rights) and = 0 otherwise.
Other	= 1 if the firm is a public company, a municipality, or a cultural, sport, religious associations and = 0 otherwise.
Industry	= 1 if the firm belongs to a specific industry, and = 0 otherwise. We have 18 different industries: Agriculture and cattle farming; Forestry and fishery; Extraction of oil and gas; minerals; Manufacturing; Electricity, gas, water; Construction; Wholesale and retail trade; Hotels and restaurants; Transport, storage and communications; Financial intermediation; Real estate activities; Public administration, defense and social security; Education; Communal and personal social services; Activities of households as employees of domestic personnel; Activities of extraterritorial organizations and bodies; Other Activities.
Bank Debt	
Outstanding Debt	total outstanding bank debt.

Table A.1: Definitions of Variables (continued)

Variable Names	Definitions
Credit Quality	
Past Non-Performance	= 1 if the firm had any repayment problems (default or delinquency) in the past 12 months, and = 0 otherwise.
Rating i	= 1 if the firm's rating is i, where i = 1 (best),. . . , 4 (worst), and = 0 otherwise.
Relationship Characteristics	
Multiple Relationships	= 1 if the firm has outstanding loans from multiple banks, and = 0 otherwise.
Rel Duration	duration of bank-firm relationship in months.
Rel Scope	= 1 if the firm has additional products (e.g., credit cards, lines of credit, discount documents, mortgages) with the bank and = 0 otherwise.
Primary Bank	= 1 if more than 50% of the firm's outstanding bank debt is originated by the bank, and = 0 otherwise.
Bank Characteristics	
Foreign Bank	= 1 if more than 50% of bank's ownership is foreign owned, = 0 otherwise.
Non-Local	= 1 if the region of the loan origination is different from a bank's headquarters, and = 0 otherwise.
Time_Entry/Foreign Owned	natural logarithm of the loan origination date minus the time since a foreign bank entry or the time since a domestic bank became foreign owned.
Cost of Deposits	average interest rate on dollar denominated saving deposits in a given month.
Market Share	a bank's total loans in the country to the total loans in the country per month.
Bank Size	natural logarithm of average total assets.
Ex Post Loan Performance	
Arrears or Default	= 1 if a loan is in arrears for more than 30 days or is downgraded to default status (rating 5), and = 0 otherwise.
Net Return on Loans	return on loans as indicated by equation (1.4.1) minus a bank's cost of deposit at loan origination.

Chapter 2

Relationship Lending and Loan Performance

2.1 Introduction

Relationship lending constitutes one of the most important comparative advantages of bank lending to firms. When engaging in relationship lending, banks gather propriety information about their customers through repeated interactions (Boot (2000)). Some theories suggest that banks use their superior knowledge to extend loans at favorable contract terms and provide firms with better access to finance (Boot and Thakor (1994), Petersen and Rajan (1995)). In contrast, alternative theories point out that relationship lending is associated with possible "hold-up" problems and extraction of rents from firms (Sharpe (1990), Rajan (1992)). While the empirical literature has mostly focused on contract terms and credit availability, far less is known about bank behavior when firms are in financial distress. This paper fills the gap by examining the effect of relationship lending on ex-post loan performance.

Understanding bank behavior when firms are hit by idiosyncratic liquidity shocks is important for practitioners and academics alike. Firms and their sources of funding when faced with liquidity shocks are vital for the economy as a whole. Moreover, insights into banks' reactions to idiosyncratic shocks to firms will help design policies to foster financial stability. While the theoretical literature has shown that informed banks provide continuation financing when firms

are in need, the empirical literature is still lagging behind. The paper uncovers a new channel of how relationship lending serves as a liquidity insurance for firms in distress.

Studies on relationship lending face several empirical challenges. First, relationship lending is not observable, being part of a bank's internal policy. Empirical studies often use the length, strength, and depth of a bank-firm relationship or the geographical distance between firms and banks as measures of relationship lending.[34] Second, the use of the relationship lending is likely correlated with firm and bank characteristics. Relationship lending centers around the acquisition of "soft" information (knowledge of the client), which plays a more important role when lending to small and opaque firms by small and decentralized banks (Stein (2002)). Ideally, one would like to observe variation in the use of relationship lending *within* banks and firms in order to eliminate biases coming from confounding effects of bank and firm characteristics. Third, one needs randomization in the use of relationship lending across banks and firms to avoid selection and composition biases.

Combining survey data on banks' lending policies with unique credit registry data from Armenia allows me to address the identification challenges and provide new insights into bank behavior when firms are in distress. First, the survey data provide a new and direct measure of the importance of relationship lending, defined as knowledge of the client, based on answers of banks. Second, the use of lending techniques varies within some banks across loan types and not others. Third, a bank-level rule helps me assign whether a loan of a firm is based on relationship or transactional lending. Since the rule is set by the bank and not publicly available, a firm is unlikely to know and thus influence the rule, creating exogenous variation (from a firm's perspective) in the use of lending techniques within a firm across banks.[35] Lastly, detailed credit registry data allow me to evaluate the effect of lending techniques within firms over time.

For the analysis, I use the bank rule to link the bank-level relationship lending measure to

[34]See among others Petersen and Rajan (1994), Berger and Udell (1995), Harhoff and Körting (1998), Machauer and Weber (1998), Degryse and Ongena (2005), Agarwal and Hauswald (2010). Elsas and Krahnen (1998) are one of the few to rely on self-evaluations of credit managers on the housebank status of a bank for a particular client.

[35]In Appendix B, I find that some manipulation is present, probably by loan officers. However, in Section 2.4.4.1 I show that identification is coming from loans further away from the threshold for which the main results continue to hold.

loan-level data from the private credit registry of Armenia for the period between January 2009 and June 2013. The credit registry data covers virtually all loans to firms during the sample period, containing detailed data on every loan such as date of origination, maturity date, contract terms and loan performance as well as data on firm characteristics such as legal status, industry, location, and banking relationships. Armenia provides an ideal setting to examine the effects of relationship lending on loan performance. By 2004, Armenia was completely privatized, leaving no government banks and only few government firms that should not distort the analysis.[36] The credit market is still at a developing stage with banks dominating the financial system.[37] Since few outside financing options exist and bank lending is of high importance, my analysis is not contaminated by other financing options for firms. While 97.7% of all registered legal entities in 2009 were small and medium-sized enterprises (SMEs), even large firms in Armenia are rather comparable to SMEs (World Bank (2014a)).[38] Most importantly, I show that stylized facts from other banking markets also hold for Armenia such as the reliance of relationship banks on "soft" information and a contract menu of high interest rates, less collateral requirements and longer maturities compared to more transaction-based banks.

To evaluate the effect of relationship lending, I distinguish between *Relationship Banks* that always rely on relationship lending and *Mixed Banks* that value relationship lending only for corporate loans and employ transactional lending based on firm fundamentals and collateral for SME loans.[39] Figure 2.1 illustrates the idea, presenting bank types, loan types and the corresponding lending techniques. Now, imagine a firm that receives an SME loan from a Relationship Bank and another SME loan from a Mixed Bank, meaning a relationship-based loan from the former and a transaction-based loan from the latter (see left arrow in Figure 2.1). Controlling for loan characteristics, I examine differences in loan performance of such a

[36] In particular, firms with private domestic ownership make up 89.9% of all firms, followed by firms with private foreign ownership (7.7%) and with state ownership (0.7%) (World Bank (2011)).

[37] The loans to GDP ratio was around 40% in 2013 compared to 96% and 198% in developed markets such as Germany and United States. According to a report by CBA (2014), banks were holding almost 90% of financial assets in 2013 and earned most of their income from lending.

[38] Relative to the European Union and the US, Armenian SMEs are much smaller with employee numbers below 100 compared to 250 and 500 employees (SME DNC (2010)).

[39] Although it might seem surprising that *Mixed Banks* do not apply relationship lending for SME loans, survey evidence by De la Torre, Martínez Pería, and Schmukler (2010) confirms that banks often offer standardized products to SMEs. Moreover, *Mixed Banks* report that they rely more on fundamentals/cash flow analysis and collateral for SME loans.

firm when lending techniques differ. In contrast, if a firm receives two corporate loans from a Relationship and Mixed Bank when both banks rely on relationship lending, I expect no differences in loan performance (right arrow in the Figure 2.1).

The effect of relationship lending on ex-post loan performance is ambiguous. Theories on financial intermediation advocate that banks mitigate asymmetric information problems by acquiring propriety information through screening and monitoring of clients in repeated interactions (see, e.g., Diamond (1984), Ramakrishnan and Thakor (1984), Allen (1990)). If relationship lending results in more efficient screening and monitoring, better ex-post loan performance should be the outcome. If banks use their knowledge to be more lenient towards a client, they might allow the client to temporarily become delinquent. Von Thadden (1995) suggests that banks might tolerate short-term bad results as long as they can extract long-term rents. The reasoning for this is that banks learn the firm's quality through monitoring, and hence will not interpret short-term bad results as signs of bad quality and prematurely terminate good projects. Similarly, Rajan (1992) suggests that informed banks offer financial flexibility to firms at the cost of a share in their profits. This paper brings these theories to the data, examining short-term loan performance and long-term effects of relationship lending.

My findings demonstrate that the *same firm* is 50% more likely to become delinquent for less or more than 90 days on a relationship-based loan relative to a transaction-based loan, indicating that Relationship Banks tolerate short-term bad results. When both banks rely on relationship lending, differences in loan performance disappear. The result continues to hold when I control for loan characteristics, bank characteristics, loan origination and firm×time fixed effects. The result is not driven by sample selection, the size of the overdue amounts, the definition of relationship lending or by relationship variables used in the previous literature.

Next, I examine the long-term effects of relationship lending. Results reveal that, albeit higher delinquency rates, Relationship Banks do not experience higher defaults or lower recovery rates when loans mature relative to Mixed Banks. Moreover, Relationship Banks earn higher returns on loans, even conditional on short-term delinquencies. In line with theory, Relationship Banks tolerate temporarily bad results as long as they can secure long-term rents. The findings uncover a new channel of how Relationship Banks provide liquidity insurance for firms that

face temporary problems.

What do Relationship Banks do differently relative to Mixed Banks that enables them to offer liquidity insurance? Empirical tests show that Relationship Banks seem to acquire information about their clients through monitoring. Not having an ex-ante riskier customer base, Relationship Banks continue to lend to customers that have been delinquent only with them, knowing their inherent quality better than Mixed Banks. Relationship Banks also allow firms to compensate temporary delinquencies with drawdowns on their credit lines or overdrafts and by adjusting interest rates and maturities but not rolling over loans more often than Mixed Banks. Overall, results give an indication of how Relationship Banks behave differently than Mixed Banks which constitutes the difference in lending technologies.

This paper makes two contributions to the literature. First, it combines a new and direct bank-level measure of relationship lending with detailed loan level data to empirically test theoretical predictions. Second, studying the effect of relationship lending on loan performance within the same firm over time, the papers offers new insights into how Relationship Banks provide liquidity insurance to firms in distress.

The paper is related to an extensive literature on relationship lending.[40] Some authors advocate that relationship lending provides better access to finance and favorable contract terms in the form of lower interest rates and less collateral requirements (see, e.g., Petersen and Rajan (1994), Berger and Udell (1995)), the more so in crisis times (see, e.g., Bolton, Freixas, Gambacorta, and Mistrulli (2013), Beck, Degryse, De Haas, and van Horen (2014a)). Others, however, indicate that banks might lock-in borrowers and raise interest rates when borrowers face transportation costs and low bank competition (Degryse and Ongena (2005)) or informational hold-up and switching costs (Ioannidou and Ongena (2010)). While previous literature mostly uses approximations of relationship lending and focuses on contract terms and credit availability, this paper combines a new and direct measure of relationship lending with credit registry data to test bank behavior when firms are in distress. A related paper by Elsas and Krahnen (1998) shows that relationship banks increase their credit volume when firms experience small rating deteriorations. This paper draws a more complete picture by revealing

[40]For a more detailed review, see Degryse, Kim, and Ongena (2009) and Kysucky and Norden (2014).

that relationship banks allow temporary delinquencies yet require long-term rents in return.

The paper also contributes to the literature on the effects of relationship lending and use of soft information on loan performance. Literature on consumer credit markets shows that relationship lending leads to lower default risk because of better screening and monitoring by banks (Puri, Rocholl, and Steffen (2013)). In mortgage markets, securitization may adversely affect screening standards of banks and result in higher default rates (Keys, Mukherjee, Seru, and Vig (2010)), while the involvement of risk managers alongside loan officers reduces default rates (Berg (2014)). Other papers show that hierarchical and physical distances decrease the use of soft information and make loan defaults more likely (Liberti and Mian (2009), Agarwal and Hauswald (2010), Skrastins and Vig (2014)). This paper shows that relationship banks treat firms differently than consumers, allowing for temporary delinquencies with future rents in sight. Some papers also explore defaults of firms but rather focus on strategic choices of firms to default on lenders with which they have personal interactions and long relationships or that are geographically and culturally close to them (Baele et al. (2014),Morales (2014),Schoar (2014)). Berlin and Mester (1999) show that banks provide insurance against credit shocks through loan rate smoothing, especially when banks rely on core deposits as a funding source. A recent paper by Li, Lu, and Srinivasan (2013) investigates contract terms during firm distress, showing that relationship banks offer similar interest rates to outside banks but keep collateral requirements low after firm distress. This paper, in contrast, shows that Relationship Banks provide liquidity insurance through financial flexibility and better access to finance.

The rest of the chapter structured as follows. Section 2.2 describes the data and provides summary statistics. Section 2.3 introduces the empirical methodology. Section 2.4 presents results and robustness tests. Section 2.5 validates the relationship lending measure. Section 2.6 concludes.

2.2 Data

2.2.1 Bank Survey Data

The first data source is the Banking Environment and Performance Survey (BEPS) II that provides information on banks' internal lending policies and organizational structures in the region of Central and Eastern Europe, Central Asia and North Africa as of 2011. Since I only have credit registry data from Armenia, I focus on Armenian banks.[41] The survey data allows me to construct a direct measure of relationship lending, relying on the answers to Questions Q6 and Q10 of 17 Armenian banks. In these answers, banks indicate the importance (frequency of use) of four lending technologies on a five-point Likert scale for SME and corporate loans separately (ranging form 1 "very unimportant" to 5 "very important"): relationship lending (knowledge of the customer), fundamental/cash flow analysis (financial information), collateral (personal assets pledged by entrepreneur), and collateral (business assets). Additionally, in Questions Q4 and Q8 banks report the number of layers that needs to be passed for loan approval if the bank grants an SME or corporate loan. In a follow-up questionnaire banks report the loan amount threshold at which they differentiate between an SME and a corporate loan.

For the main empirical analysis, I use the survey answers on the absolute importance of relationship lending. In Section 2.4.4, I also experiment with a measure of the relative importance of relationship lending, that is relative to other lending techniques. In contrast to previous literature, I do not approximate relationship lending with the strength, length, depth or distance of bank-firm relationships but directly observe the self-reported importance of relationship lending of each bank for two loan types.[42] The survey specifies relationship lending as knowledge of the client, which is similar to Boot (2000)'s definition of relationship lending as the acquisition of propriety information over multiple interactions with a customer. Although

[41] The BEPS II survey was jointly undertaken by the European Bank for Reconstruction and Development (EBRD) and European Banking Center (EBC) at Tilburg University in 2011. CEOs of more than 400 banks in the region of Central and Eastern Europe, Central Asia and North Africa were interviewed by a specialized team of senior financial consultants with considerable banking experience on topics such as activities, funding and risk management, lending technologies, competitive environment, the influence of foreign parent banks and perceptions of legal and regulatory systems.

[42] Based on the same survey, Beck et al. (2014a) construct the share of relationship banks in the vicinity of firms and test how it affects access to finance for these firms over the business cycle.

banks report their lending technologies only in 2011, it should not pose a problem, since Fahlenbrach et al. (2012) confirm that bank business models stay relatively constant over time. Still, I conduct several robustness tests in Section 2.4.4.1 and show that the relationship lending measure captures the use of "soft" information when pricing loans in Section 2.5. Moreover, Beck, Degryse, De Haas, and van Horen (2014a) provide cross-country evidence for the validity of the relationship lending measure.

Based on banks' survey answers I define 9 *Relationship Banks* that report relationship lending (knowledge of the client) to be "very important" for both SME and corporate loans, and 6 *Mixed Banks* that find relationship lending "very important" for corporate loans only but less for SMEs. Figure 2.1 shows the two bank types, loan types and the corresponding lending techniques.[43] It might seem surprising that Mixed Banks do not rely on relationship lending but rather on transactional lending for SMEs. In fact, these banks mostly report firm fundamentals, cash flow analysis and collateral to be more important for giving out SME loans. Moreover, 97.7% of all registered legal entities in 2009 were SMEs, implying that even large firms in Armenia are rather comparable to SMEs in more developed countries (World Bank (2014a)). SME loans are thus relatively small loans for which standardization of the lending process is more likely. Survey evidence by De la Torre et al. (2010) shows that banks offer standardized products, sometimes combined with tailor-made features, to a group of SMEs with similar needs such as adjusting the credit line to the business cycle of a specific industry.

The second important information are the bank-specific definition criteria for SME and corporate customers. Most banks base the customer type on the loan amount.[44] Using the loan amount threshold, I can assign a loan to be a SME or corporate loan and the corresponding lending technique the bank uses for the loan type (see Figure 2.1). The loan amount thresholds differ across banks in size and currency and are not publicly available, creating exogenous variation (from a firm's perspective) in the assignment of a firm's loan to a different loan type

[43] Two banks drop from the sample because one bank grants SME loans only and the other one never considers relationship lending as "very important" for both loan types.

[44] Although there exists a law on "State Support to SMEs" that specifies annual assets and turnover as well as the number of employees for different firm categories, banks use their own rules of thumb to classify a firm as SME or corporate firm. Discussions with bank employees, ACRA and CBA employees, confirmed that banks classify firms mostly by loan amount. An IFC survey also shows that the loan amount is a good proxy for firm size (IFC (2013)).

and hence to the corresponding lending technique.[45] Such bank policies can still be subject to manipulation by loan officers. In Appendix B, I find that some manipulation is present. In Section 2.4.4.1, however, I show that identification is coming from loans further away from the loan amount threshold for which main results continue to hold.

2.2.2 Credit Registry Data

The second unique dataset is the Armenian private credit registry from the Armenian Credit Reporting Agency (ACRA) that allows me to assess the effect of relationship lending on loan outcomes.[46] Founded in 2004, the credit registry covers all loans to firms (without a loan amount restriction) from 21 Armenian banks between January 2009 and June 2013 on a semiannual basis.[47] For each loan, I have information on the origination and maturity dates, contract terms, ex-post loan performance, location of loan issue, and economic sector of a loan. For each firm, I have information on the legal status, industry, physical location, and banking relationships. The private credit registry is used by the Central Bank of Armenia (CBA) for the analysis of the banking sector as well as for supervision purposes. Since all banks are members of the private credit registry and have paid a flat rate membership fee, they mostly rely on data from the private credit registry, which is more complete, for screening and monitoring purposes of firms. In almost 99% of the cases banks inquire information in a standardized format that covers information on all current and past loans of a firm dating back 5 years without revealing the other bank's identity.[48]

[45] Liberti and Mian (2009) also use 19 bank rule variables that determine the hierarchical level to which a firm is sent in order to investigate the effect hierarchical distance on information production.

[46] Armenia has also a public credit registry managed by the Central Bank of Armenia (CBA), which, however, covers only large loans above around 3.165 million US dollars (1.5 million Armenian Drams) that do not receive the highest credit rating, while the private credit registry covers virtually all loans to firms. Based on the "Procedure for Creation of Information System of Creditworthiness of Customers of Banks, Credit Organizations, Branches of Foreign Banks operating in the Republic of Armenia, that is of Credit Registry, and Procedure for Participation in Credit Registry", banks are obliged to report to all credit registries about all their loans within 3 business days. Discussions with ACRA and CBA stuff confirm that banks report on their loan portfolio almost on a daily basis.

[47] I received access to the credit registry through a collaboration of the CBA, the EFSE Development Facility and Tilburg University.

[48] Typical information includes contract terms, ex-post loan performance, firm characteristics and some information on borrower-affiliated parties. Information in a non-standardized format additionally covers detailed information on firm owners, participants, and guarantors but requires additional consent from involved parties and is only available against a much higher fee than the usual membership fee and is therefore only rarely requested.

The Central Bank of Armenia has merged the BEPS II survey data with the credit registry data, keeping the bank and firm names anonymous. Using the loan amount thresholds, I assign each loan of a firm to be SME or corporate. Therefore, I convert data on bank level to loan level. The final sample consists of 15 banks that report the importance of use of the relationship lending, and account, on average, for 84% (87%) of all banks' credit contracts in terms of value (number). The credit registry covers different contract types such as loans, credit lines, factoring, leasing, guarantees, letters of credit, overdrafts, repurchase agreements, and swaps. I focus on loans and refer to them as "standard credit contracts" to have a more homogenous and most representative set of loans.[49] These contracts account for 73% (70%) of all credit contracts in terms of value (number). Most loans are issued either in Armenian Drams (AMD) or US dollars (49% and 48%, respectively) with only few loans in euros, British Pounds and Russian Rubles. I convert all loan amounts in US dollars based on the monthly exchange rate of the Armenian Central Bank.[50] To ensure the use of timely information, I only study "new loans" that have been issued between January 2009 and June 2013.[51] Few loans that are not rated or have a written-off status at the date of loan issue and loans that have a zero interest rate (but possibly have non-zero fees) are dropped. The resulting dataset consists of 53,780 loan-time observations of 19,332 loans to 6,649 firms with an average loan spell of almost 3 years.

Differences in firm composition might influence the effect of relationship lending on loan performance. For example, high-risk firms might self-select themselves to Relationship Banks, expecting that these banks will grant them more freedom to become delinquent on their loans. Therefore, I focus on firms that received loans from both bank types, Mixed and Relationship Banks for the main analysis. Since sample selection is not random, inferences based on such a selected sample do not necessarily apply to the total sample but offer a cleaner identification not contaminated by firm selection and composition effects. The sub-sample consists of 10,656 loan-time observations of 4,441 loans to 621 firms and constitutes nearly 40% of the total lending

[49] I use other credit contracts to calculate the exposure of firms to banks as well as relationship characteristics.

[50] Note that base interest rates might differ across loans of different currencies. Since I focus on loan performance in Section 2.4.1 and the variation in interest rates in Section 2.5, it should not affect the results. Still, I control for differences in currencies and conduct robustness tests for loans of different currencies.

[51] I exclude loans that have been originated before January 2009 and stay in the system because of past non-performance.

amount of the entire sample. To account for time-varying firm characteristics, I also experiment with a sample of 3,790 loan-time observations of 1,952 loans to 318 firms that received loans from both bank types in the same period.[52]

2.2.3 Descriptive Statistics

In order to gain some insights into banks' survey responses as well as other characteristics of Armenian banks, Table 2.1 reports bank types (Relationship vs. Mixed Banks), survey responses on the importance of relationship lending and the average number of hierarchical layers for loan approval by loan type as well as bank level summary statistics on the average loan amount threshold in US dollars, the market share in terms of loan number and loan value, the average borrower size based on the total borrowing amount across all banks in US dollars, bank size based on total assets as of 2009 and foreign ownership that equals one if more that 50% of equity is foreign-owned. Since bank names remain anonymous, I rely on bank characteristics from the BEPS II survey and reported by the Central Bank of Armenia. A first look reveals that bank characteristics vary across the two bank types, Mixed and Relationship Banks. Unreported results suggest that no specific bank characteristic is correlated with the importance of relationship lending in Armenia, confirming that the two bank types are not driven by other bank characteristics. The only difference between the two groups is generally lower loan amount thresholds for Relationship Banks.[53] The differences in loan amount thresholds rather facilitate the analysis as it allows for more variation in loan types within one firm across the two bank types.

Table 2.2 shows ex-post loan performance measures, loan characteristics, relationship characteristics, and firm characteristics by bank type for the total sample of 19,332 loans to 6,649 firms and for the sub-sample of 4,441 loans to 621 firms that borrow from both bank types. Definitions of the variables can be found in Table A.1 of Appendix A. To capture temporary loan delinquencies, I mainly rely on performance measures "Non-Performance 0-90 days" that is equal to one if a loan goes overdue for less than 90 days and zero otherwise as well as "Non-

[52]Still, in Section 2.4.4.1, I confirm that main results continue to hold for the total sample as well as the sample of firms that exclusively lend from *Relationship* or *Mixed Banks*.

[53]In unreported regressions results of the determinants of loan amount thresholds, I find that larger banks are more likely to have higher loan amount thresholds.

Performance 90 days" that is equal to one if a loan goes overdue for more than 90 days and zero otherwise. Other measures of loan performance indicate if a loan has any overdue principal or interest payments (Non-Performance), goes overdue for more than 180 days (Non-Performance 180 days), or is completely written off/lost (Default).

Descriptive statistics and graphical analysis on loan performance offer first evidence that Relationship Banks temporarily experience higher delinquency rates without more loan defaults. For the total sample, loan delinquencies are significantly higher by 0.8% to 2.6% for Relationship Banks compared to Mixed Banks. For the sub-sample, the differences in loan delinquencies for less and over 90 or 180 days increase, while the differences in loan defaults vanish. Figure 2.2 shows average delinquency rates over 90 days over time for the sub-sample by bank type (Relationship vs. Mixed Banks) and loan type (Figure 2.2a for SME loans and Figure 2.2b for corporate loans). Delinquency rates over 90 days are always higher for Relationship Banks compared to Mixed Banks for SME loans (Figure 2.2a), the case when Relationship Banks rely on relationship lending and Mixed Banks on transaction-based lending. In case of corporate loans (Figure 2.2b), when both employ relationship lending, no consistent differences in delinquency rates are present. The figures confirm that relationship lending leads to temporarily higher delinquencies independent of time.

Apart from firm compositional and selection biases, differences in contract terms between Relationship and Mixed Banks might influence loan performance. For the total sample, Relationship Banks give out smaller loans ($181,000 versus $225,000) at slightly higher interest rates (15.4% versus 15.1%) that are less likely to be collateralized (82% versus 88%) or guaranteed (5% versus 10%) but are granted at almost 3 months longer maturities (36 months versus 33 months) without differences in credit ratings compared to Mixed Banks. The findings are consistent with Beck, Ioannidou, and Schaefer (2014b) who show that domestic banks are more likely to give out unsecured loans at higher interest rates and longer maturities than foreign banks. Except for collateral and guarantees, most of the differences in contract terms disappear for the sub-sample, suggesting that differences in loan characteristics should not be driving differences in loan performance.

If contract terms do not significantly differ across bank types within the same firm, then I can isolate the effect of relationship lending on loan performance. Keys, Mukherjee, Seru,

and Vig (2010) claim that when a firm applies for a loan, banks collect "hard" information (e.g., credit rating) and "soft" information (e.g., knowledge of the client or future repayment ability). When studying loan performance around a FICO score threshold, they assume that after controlling for hard information any differences in default rates on either side of the threshold should be coming from the effect of securitization on the use of soft information. Similarly, the results above show that there are no significant differences in loan characteristics for the sub-sample. Still, in the main analysis, I control for observable hard information variables (loan characteristics) and unobservable soft information (firm and firm×time fixed effects). The relationship lending measure thus should capture some of the unobservable soft information "extracted" from the firm or firm×time fixed effects and the residual. In Section 2.5, I show more formally that relationship lending is indeed associated with higher use of soft information.

Differences in other loan characteristics (location, industry and currency), relationship characteristics (bank-firm relationship in months, scope, primary bank, number of relationships) and other firm characteristics (firm location, industry, legal type) do not completely disappear for the total sample compared to the sub-sample. Since the main analysis focuses on variation within a firm, most of these differences will not play a role. Moreover, I control for these characteristics in different specifications of the model and in robustness tests.

2.3 Empirical Methodology

In order to identify the effect of relationship lending on ex-post loan performance, I distinguish between *Relationship Banks* that always rely on relationship lending and *Mixed Banks* that value relationship lending only for corporate loans and employ transactional lending based on firm fundamentals and collateral for SME loans. Figure 2.1 illustrates the idea, presenting bank types, loan types and corresponding lending techniques. Mixed Banks are the treatment group since they rely on transactional and relationship lending for SME and corporate loans respectively. Relationship Banks are the control group, always employing relationship lending. Using the loan amount threshold, I assign SME and corporate loan types to each loan of a firm. The loan type then determines the lending technique for each loan of a firm.

Now, imagine a firm that receives an SME loan from a Relationship Bank and another SME loan from a Mixed Bank. Controlling for loan characteristics, I examine differences in loan performance of such a firm when lending techniques differ (relationship versus transactional lending, the left arrow in Figure 2.1). In contrast, if the firm receives two corporate loans from a Relationship and Mixed Bank when both banks rely on relationship lending, I expect no differences in loan performance (right arrow in Figure 2.1). Empirically, I estimate the following model:

$$\begin{aligned} \text{Loan Performance}_{ijkt} = {} & \beta_1 \text{Corporate Loan}_{ijk} \\ & + \beta_2 Mixed\,Bank_k \\ & + \beta_3 \text{Corporate Loan}_{ijk} \times \text{Mixed Bank}_k \\ & + \theta\text{'Controls}_{ijk} + \alpha_j + \gamma_t + \varepsilon_{ijkt}, \end{aligned} \tag{2.3.1}$$

where i, j, k, t index loans, firms, banks, and time (semiannually). Notice that time stands for the period each loan is observed in and not only the loan origination period. $\text{Loan Performance}_{ijkt}$ equals 1 if a loan becomes delinquent for (less) more than 90 days in a given half year. $\text{Corporate Loan}_{ijk}$ equals one if a firm receives a corporate loan from a Relationship Bank. $Mixed\,Bank_k$ equals one if a firm receives an SME loan from a Mixed Bank. $\text{Corporate Loan}_{ijk} \times \text{Mixed Bank}_k$ equals one if a firm receives a corporate loan from a Mixed Bank. The reference group are SME loans of Relationship Banks.

The coefficient β_2 is the main coefficient of interest that gauges the differences in performance for transaction-based versus relationship-based loans for the *same firm* (SME loans of Mixed vs. Relationship Banks). A negative coefficient means that transactional lending leads to less delinquencies, implying that relationship lending leads to more delinquencies of SME loans. A positive coefficient, in turn, indicates that transactional lending results in more delinquencies and thus less delinquencies of SME loans for relationship lending. The difference between the coefficients β_1 and β_3 measures the difference in performance when both banks employ relationship lending (corporate loans of Mixed vs. Relationship Banks).

Controls_{ijk} consists of loan, firm and bank characteristics. Loan characteristics include contract terms such as the credit rating at loan initiation ($\text{Credit Rating}_{ijk}$), loan interest

rate (Interest Rate$_{ijk}$), the natural logarithm of one plus the loan amount (Loan Amount$_{ijk}$), two dummy variables that indicate whether the loan is collateralized (Collateral$_{ijk}$) or has a guarantee (Guarantee$_{ijk}$), and the natural logarithm of one plus the loan maturity in months (Loan Maturity$_{ijk}$). Although results in Section 2.2.3 show that most contract terms do not differ across Relationship and Mixed Banks, I still control for them in the regressions. I assume that contract terms are hard information variables that explain the performance of a loan. Controlling for hard information variables within the same firm ensures that the effect on loan performance is coming from differences in the importance of relationship lending and thus the use of unobservable soft information (see, e.g., Keys et al. (2010)).[54] If omitted variable bias is driving the results it will affect Relationship and Mixed Banks in the same way since contract terms do not differ across the two bank types.

To account for observable firm characteristics, I include dummy variables that indicate whether a firm is located in the capital Yerevan, whether it is associated with the trade or other fields of service industries and whether the firm is a private firm. Since bank characteristics might influence bank behavior, I use indicators of bank size, ownership and average hierarchy for loan approval. Finally, I also include dummy variables that indicate loan location, industry and currency.

The firm fixed effects, α_j, eliminate firm heterogeneity and compositional biases, comparing differences in loan performance within the same firm across different banks. Time of loan origination fixed effects, γ_t, are included to control for the timing of loan origination. In some specifications, I also use firm×time fixed effects (α_{jt}), where time stands for the current period, to account for changes in firm characteristics over time. Standard errors are always clustered at firm level to control for possible correlations in the residuals across observations of the same firm.[55] In the main regressions, I rely on a linear probability model since logit models with a large number of fixed effects suffer from the "incidental parameter problem".[56] In section

[54] In Section 2.5, I confirm that relationship lending is indeed associated with higher use of soft information in loan pricing.

[55] In Section 2.4.4.1, I show that results survive clusters at bank level and double clustering at bank×time and firm levels.

[56] Because of the large number of fixed effects in the model relative to the smaller number of periods for which a borrower is observed, a non-linear model could give inconsistent estimates; this is known as the "incidental parameter problem" (see, for example, discussion in Cameron and Trivedi (2005, pp. 726-727)).

2.4.4.1, I confirm the main results using a logit model.

2.4 Relationship Lending and Loan Performance

In this section, I first show that relationship lending results in temporarily higher delinquency rates. Second, I reveal that, given previous delinquencies, Relationship Banks are not worse off in the long run, are able to extract rents and continue to lend to firms afterwards. Third, I examine the behavior of Relationship Banks that enables them to offer liquidity insurance. Finally, I present several alternative explanations and robustness tests.

2.4.1 Relationship Lending and Loan Delinquencies

In Figure 2.2, we have already observed that firms are more likely to become delinquent on relationship-based relative to transaction-based loans. Tables 2.3 and 2.4 document more formally that temporary delinquencies are less likely for SME loans of *Mixed Banks* relative to SME loans of *Relationship Banks* based on the coefficient of Mixed Bank$_k$. Table 2.3 uses delinquencies for less than 90 days and Table 2.4 for delinquencies over 90 days as dependent variables for the sub-sample of 621 firms that receive 4,441 loans from both Relationship and Mixed Banks**.** Columns (2) of both tables suggest that the *same firm* is by 2.5 (1.6) percentage points more likely to become delinquent for less (more) than 90 days on a relationship-based relative to a transaction-based loan. Given an average delinquency rate for less (more) than 90 days of 4.9% (3%) for the sub-sample, the results imply that a firm is by around 50% more likely to become temporary delinquent on a relationship-based loan, an economically meaningful effect.[57]

In all specifications, I control for loan contract characteristics to isolate the effect of relationship lending on loan performance from observable hard information variables that explain loan performance. To show the consistency of results, I run a pooled regression with firm characteristics in Column (1) to control for hard information variables on borrower level; in Columns (2)-(5) I use firm fixed effects to account for any unobservable firm characteristics and in Column (6), I add firm×time fixed effects to additionally control for changes in firm char-

[57]Results continue to hold for a logit model with similar magnitudes of the odds ratios of the main effect (see Section 2.4.4.1 and Table 2.10).

acteristics over time. In Columns (3)-(5), apart from firm fixed effects, I add loan origination fixed effects, bank characteristics and other loan characteristics.

Comparing Columns (1) and (2) in Tables 2.3 and 2.4, it becomes evident that adding firm fixed effects does not change the main coefficient β_2, suggesting that the effect is not driven by differences in firm characteristics but rather differences in lending technologies (relationship lending for SME loans of Relationship Banks and transactional lending for SME loans of Mixed Banks). Even when firm×time are added in Column (6), the main effect remains significant for delinquencies below 90 days but loses it significance for delinquencies above 90 days.[58] While loan origination fixed effects in Column (3) and other loan characteristics in Column (5) do not change the magnitude of the main coefficient, adding bank characteristics in Column (4) increases the effect to 3.2 (1.9) percentage points in Table 2.3 (2.4). Overall, the main coefficient remains highly significant and similar in size throughout all specifications.

To further strengthen my results, I test for differences between coefficients of corporate loans of Relationship and Mixed Banks, the case when both rely on relationship lending. The p-values at the bottom of Tables 2.3 and 2.4 are never highly statistically significant, confirming that only differences in lending techniques affect loan performance. Additionally, the coefficient of the $\text{Corporate Loan}_{ijkt} \times \text{Mixed Bank}_k$ is often positive and significant in Table 2.4, revealing that also within Mixed Banks relationship lending leads to higher delinquencies.

Turning to loan characteristics, I find that initial credit ratings have the highest economic effect, being a natural predictor of future loan performance. A one unit increase in the credit rating is associated with a 18 to 39 percentage points lower delinquency rate. Loans that have high interest rates, are collateralized and have longer maturities are more likely to become delinquent. Higher loan amounts are generally associated with higher delinquencies rates over 90 days but seem to lead to less delinquencies for loans within the same firm. As for other control variables, higher hierarchies for loan approval have a slightly significant positive effect on delinquencies for less than 90 days, while loans in other service industries experience lower delinquency rates. The other variables do not exhibit a significant effect on loan performance.

All in all, results indicate that relationship lending leads to temporarily higher delinquencies

[58] For the logit model, however, the effect is significant for both delinquency measures.

of less or more than 90 days. Von Thadden (1995) and Rajan (1992) suggest that banks might tolerate short-term bad results and thus offer financial flexibility as long as they can extract long-term rents. Banks learn a firm's quality through monitoring during the loan spell and therefore do not consider bad short-term results as a sign of bad quality and prematurely terminate good projects. In the next section, I explore what happens in the long run.

2.4.2 Relationship Lending and Long-Term Effects

In this section, I examine long-term effects by looking at defaults, recovery rates, and losses as well as return on loans at the end of the loan spell. If Von Thadden (1995)'s and Rajan (1992)'s theories hold, I should not find differences in loan performance at the end of the loan spell, despite temporary higher delinquencies for relationship-based loans. Knowing the quality of the firms, Relationship Banks will allow temporary delays in loan repayment for firms that are not likely to default at the end of the loan spell. Still, Relationship Banks might extract higher rents through, for example, higher return on loans in order to be compensated for such temporary delinquencies.

Descriptive statistics in Table 2.2 have already shown that, for the sub-sample, Relationship Banks have significantly higher delinquency rates but not higher default rates (in loss/written-off status) compared to Mixed Banks. In Table 2.5, I examine only SME loans of Relationship and Mixed Banks, the case when one relies on relationship lending and the other on transactional lending. To account for right-censoring, I select SME loans that are observed until maturity and condition them to have been delinquent during the loan spell. Since many observations are lost, I present results for both the total and the sub-sample, also because the effect of higher delinquencies of Relationship Banks continues to hold for the total sample (see Section 2.4.4.1). To measure long-term performance, I calculate loan defaults that equal one if a loan is in loss or written-off status at the end of the loan spell, recovery rates that equal one if a loan has had overdue payments during the loan spell but did not default at maturity and losses defined as written-off amounts relative to the contract amount. As I do not have a direct measure of written-off amounts, I use the overdue principal (plus interest rate amount) given delinquencies. Since I am interested in looking at long-term effects for the two lending

technologies, I focus on cross-sectional tests for differences in long-term performance.

Panel A of Table 2.5 reports univariate tests of default rates, recovery rates, and losses at the end of the loan spell for SME loans that have been delinquent for less than 90 days. Although SME loans of Relationship Banks seem to default more often after delinquencies for the total sample, the difference is not highly statistically significant and disappears for the sub-sample. Recovery rates always remain insignificant for the total sample and sub-sample. Interestingly, the losses in percentage of the contract amount are smaller for Relationship than Mixed Banks, indicating that despite higher delinquency rates Relationship Banks are not losing more on average.

Panel B repeats the analysis for SME loans that have been delinquent for more than 90 days. While loan defaults are higher and recovery rates lower for the total sample at 10% significance level only, differences again disappear for the sub-sample. Losses given delinquencies over 90 days do not differ for both samples. However, the amounts lost are much higher for delinquencies over 90 days than for short-term delinquencies below 90 days, since the former once are more likely to turn into actual defaults. Overall, the results indicate that relationship lending does not lead to worse long-term performance at the end of the loan spell, despite higher temporary delinquencies. In unreported regression results, I confirm that results continue to hold when I use regression analysis.

Long-term performance at the end of the loan spell might be correlated if firms default across all banks. For this reason, I also calculate return on loans a bank gains at the end of the loan spell or until the loan is observed. Table 2.2.3 has already revealed that Relationship Banks charge higher interest rates for the total sample, suggesting that Relationship Banks require higher rates for the financial flexibility they offer. Following Haselmann, Schoenherr, and Vig (2013), I calculate the return on a loan (ROL) by bank j to firm i for the entire loan spell:

$$\text{ROL}_{ij} = \sum_{t=1}^{T} \frac{\text{Loan Balance}_{ijt}}{\sum_{t=1}^{T} \text{Loan Balance}_{ijt}} \left[(1 - \mathbb{1}_{\{NPL=1\}}) r_{ijt} + \mathbb{1}_{\{NPL=1\}} \text{Loss}_{ijt} \right], \tag{2.4.1}$$

where the first term stands for the ratio of the outstanding loan amount from bank j to firm

i at the beginning of period t ($\text{Loan Balance}_{ijt}$) to the sum of the outstanding loan amounts over the loan spell ($\sum_{t=1}^{T} \text{Loan Balance}_{ijt}$). The indicator function ($\mathbb{1}_{\{NPL=1\}}$) equals one when a loan has overdue amounts between t and t+1, r_{ijt} is the interest rate charged by bank j to firm i and Loss_{ijt} is the loss of the bank, which is defined as the (negative) of the written off amount over the contract amount. The weights ensure that returns or defaults receive more weight at the beginning of the loan spell and less weight at the end of the loan spell, when most of the loan has been repaid.[59] Since I do not observe the actual written off amounts, losses are defined as the overdue principal (plus overdue interest) amount in case a loan is in a loss/written-off status or delinquent over the contract amount.[60]

Table 2.6 shows univariate tests and regression estimates of ROLs of SME loans for the total sample and the sub-sample. Unconditional ROLs of SME loans are significantly higher for Relationship Banks relative to Mixed Banks for the total sample but not for the sub-sample. While ROLs given delinquencies below 90 days are by 2.5 percentage points higher for Relationship Banks and statistically significant for both samples, ROLs given delinquencies over 90 days are not significantly different for both samples. The last panel presents pooled regression results of the ROL measure on a Mixed Bank dummy for the total sample and sub-sample of SME loans, unconditionally and conditional on delinquencies below and over 90 days. Results are similar to univariate tests with significantly higher ROLs for the total sample for all SME loans and given delinquencies below 90 days as well as for the sub-sample for delinquencies below 90 days. Although differences in ROLs between Relationship and Mixed Banks are not always statistically significant, they are always larger in size for Relationship Banks.

In sum, results in this section are in line with theoretical predictions. Relationship Banks allow their customers to become temporarily delinquent yet are able to extract long-term rents, not facing higher defaults or losses but ofter higher return on loans when loans mature. These findings imply that relationship lending provides a liquidity insurance for firms in distress, offering greater financial flexibility without incurring higher losses and often earning higher

[59] Haselmann, Schoenherr, and Vig (2013) additionally discount the weights to account for the time value of money. Unreported results using discounted ROLs confirm that the findings are not affected.

[60] In unreported results, I also use the overdue principal (plus overdue interest) amount over the contract amount only in case a loan is in a loss/written-off status and zero otherwise. This measure yields virtually the same results.

rents.

2.4.3 Relationship Lending and Bank Behavior

Results in the previous sections suggest that Relationship Banks must have some superior technology that allows them to be more flexible and forgiving to their customers. What do Relationship Banks do differently relative to Mixed Banks that enables them to offer liquidity insurance? This section tries to answer this question by examining the distribution of borrower's ex-ante riskiness, loan rescheduling and drawdowns on credit lines and overdrafts given delinquencies as well as monitoring behavior of banks.

Differences in borrowers's ex-ante riskiness might drive the behavior of Relationship Banks when firms are in distress. Relationship Banks might be able to offer liquidity insurance to customers simply because they have an overall better customer base than Mixed Banks. In turn, Relationship Banks might also have a worse customer base that makes them adjust their lending behavior. Therefore, Panel A of Table 2.7 compares ex-ante non-performance with the bank a firm borrows from (Past NPL Bank), with any bank (Past NPL Any Bank) including the former, and with other banks (Past NPL Other Banks) between Relationship and Mixed Banks on bank-borrower-time and bank-borrower level for the total sample of firms. Results reveal that customers of Relationship Banks have a worse performance history with their Relationship Bank relative to customers of Mixed Banks with their Mixed Bank. This indicates that Relationship Banks are more likely to continue to lend to firms despite past non-performance, which is another proof for the insurance function of relationship lending. There are, however, no statistically significant differences in ex-ante loan performance with any bank or other banks for the customer base of both bank types. This implies that Relationship Banks must have some additional unobservable information about their clients that allows them to be more lenient and not make losses long term.

Another way to explore the behavior of Relationship Banks is to look at loan rescheduling and renegotiations. Are Relationship Banks more likely to rollover loans and renegotiate contract terms when firms are in trouble? Panel B and C of Table 2.7 show the average percentage of rollover loans within one-month, two-month, and three-month windows around the maturity

of a loan with less or more than 90 days delinquencies.[61] Rollovers occur, in general, in 1.2% to 2.7% of the cases for both samples but there are no highly significant differences between Relationship and Mixed Banks despite higher delinquencies among SME loans of Relationship Banks. This suggests that Relationship Banks do not evade defaults by issuing new loans.

Panel B and C also report the average percentage of increases and decreases in the interest rate, loan amount and loan maturity for the total sample and sub-sample given delinquencies below or above 90 days, respectively. In particular, changes in contract terms are measured on loan-time level at the period and up to three periods after delinquencies and then collapsed to loan level. Relationship Banks are more likely to increase maturities upon delinquencies than Mixed Banks for both samples. Although lengthening maturities might suggest that Relationship Banks avoid delinquencies, Table 2.7 also shows that they shorten maturities almost as often. They are also more likely to increase or decrease interest rates for the total sample. This indicates that Relationship Banks are in general more active in adjusting contract terms upon delinquencies which might help them alleviate the liquidity shortages of firms.

Since firms often have other contracts such as credit lines and overdrafts with banks, Relationship Banks might allow firms to draw down on them once in distress, mitigating future defaults. Panel D of Table 2.7 reports whether firms with SME loans and credit lines or overdrafts with the same bank draw down on them upon delinquencies or up to three periods later on loan-time level and then collapsed to loan level. Since these restrictions reduce the sample size, I only look at the total sample of SME loans and additionally report drawdowns given both delinquency types (less or more than 90 days). Clients of Relationship Banks draw down on credit lines or overdrafts in around 37 to 42 percent of delinquency cases of over or/and below 90 days, while clients of Mixed Banks only in around 23 to 32 percent. The differences are large in magnitude but not that significant due to the low number of observations. Conditioning on any overdue payments or overdue days in the last row yields more observations and a similar difference of 10% that is highly significant. Confirming Von Thadden (1995)'s predictions, Relationship Banks seem to combine normal contracts with credit lines and overdrafts to allow

[61] As I do not know for sure whether a loan has been rolled over, I declare a loan of a firm to be a rollover loan if the loan has been non-performing and a new loan has been issued within one, two, or three months before or after the the non-performing loan with the same bank.

firms to draw down on them in distressed times.

Since Relationship Banks do not have an ex-ante different customer base, they must be screening or monitoring firms to acquire superior knowledge about their clients. Although I do not directly observe the behavior of banks, I have data on the use of the credit registry by banks to inquire information about firms for loan granting and loan monitoring purposes between June 2012 and June 2013. Since the data just offers the numbers of inquiries on firm level, I focus on firms that received loans either only from Relationship or Mixed Banks to be able to disentangle the effect for these two bank types. Table 2.8 reports regression results for the number of inquiries for loan monitoring in Columns (1)-(3) and loan granting purposes in Columns (4)-(6) for a sample of 2,737 firms that exclusively borrow either from Relationship or Mixed Banks and for which data on banks' use of the credit registry exists. The main independent variable is the *Mixed Bank Firm* dummy that tests for differences in the use of credit registry between the firms that received loans just from Mixed relative to Relationship Banks. To control for firm heterogeneity, I add firm size measured as the average total outstanding debt across banks, past non-performance measures with the same bank, any bank and other banks as presented above as well as firm location, industry and ownership fixed effects. The coefficient of *Mixed Bank Firm* is negative and significant in Columns (1)-(3), suggesting that *Relationship Banks* are by around 26 to 33 percentage points more likely to inquire the credit registry to monitor firms. For loan granting inquiries of the credit registry, the *Mixed Bank Firm* dummy is never significant and small in size. These results suggest that Relationship Banks are more likely to monitor firms but do not inquire additional information when granting loans.

In line with Von Thadden (1995)'s theoretical predictions, Relationship Banks seem to acquire information about their clients through monitoring. Not having an ex-ante riskier customer base, Relationship Banks continue to lend to customers that have been delinquent only with them, knowing their inherent quality better than Mixed Banks. Relationship Banks also allow firms to compensate for delinquencies with drawdowns on credit lines or overdrafts and by adjusting interest rates and maturities but not rolling over loans more often than Mixed Banks. Overall, results give an indication of how Relationship Banks behave differently than

Mixed Banks which constitutes the difference in lending technologies. Section 2.5 additionally confirms that relationship lending is based on use of soft information about clients.

2.4.4 Alternative Explanations

So far, we have established that Relationship Banks allow their customers to temporary become delinquent but do not suffer from higher defaults and often higher returns when loans mature. In this section, I explore whether the effect of relationship lending on delinquencies below 90 days survives alternative explanations.[62] For all tests in Tables 2.9 and 2.10, I use specifications of Table 2.3 in Columns (1) and (2) with firm characteristics and with firm fixed effects. To consume space, Table 2.10 reports only the main coefficient of $\text{SME Loan}_{ijk} \times \text{Mixed Bank}_k$ that measures the effect of low importance of relationship lending on loan performance as well as its standard error and the number of observations.

First, I investigate whether firms are more likely to delay repayments with Relationship Banks because the delayed repayments are small in size. To this extent, I create three ratios that capture the relative size of delayed repayments: all measures use the overdue principal and interest rate amount in the numerator and set it relative to the total loan amount (NPL Amount over Contract Amount), the total outstanding firm debt with the particular bank in a period (NPL Amount over Bank Exposure) or the total outstanding firm debt with all banks in a period (NPL Amount over Total Exposure). For each dependent variable, I use firm characteristics or firm fixed effects to control for unobserved firm heterogeneity and restrict the sample to loans that have been delinquent for less than 90 days. Results in Table 2.9 show that there is never a statistically significant difference of relative overdue amounts for SME loans of Relationship and Mixed Banks, even within the same firm. This result confirms that higher delinquencies with Relationship Banks cannot be explained by smaller delayed repayment amounts.

Second, relationship lending might not be the most important lending technology for a bank such that a 5 for "very important" (in the survey) might mean something else for one bank compared to another bank. Therefore, I construct a measure of the importance of relationship lending relative to the importance of other lending technologies (fundamental, private and

[62] In unreported results, I confirm that alternative explanations also hold when I use logit models instead or delinquencies over 90 days as the dependent variable.

business collateral) by loan type. For all Mixed Banks the relative importance of relationship lending is always higher for corporate compared to SME loans. The importance of relationship lending stays constant across loan types for only 4 out of 9 Relationship Banks and is always higher compared to Mixed Banks. Row "Alternative relationship lending" of Table 2.10 reveals that for the reduced sample of banks the main result continues to hold, although somewhat reduced in magnitude and significance. Still, the relationship lending measure might simply capture relationship variables used in the previous literature. In Row "Relationship variables", I therefore add relationship variables to the main regressions such as the natural logarithm of one plus the bank-firm relationship in months, a dummy variable that equals one if a firm has more than 50% of its outstanding debt with a bank, and a dummy variable that equals one if a firm has other products with the bank (e.g., credit lines, factoring, guarantees). Although, in unreported results, the dummy variable for the main bank of a firm significantly lowers delinquencies below 90 days, the negative effect of transactional lending on loan performance is not affected. This finding confirms that the relationship lending measure captures something beyond typical relationship lending proxies, most likely the use of firms' unobserved soft information.

Third, the main results in Tables 3 and 4 rely on a sub-sample of firms that receive loans from both Relationship and Mixed Banks. Since results might suffer from a selection bias, in Row "Full/Opposite Sample" of Table 2.10 I estimate specification (2) of Table 2.3 with firm characteristics for the full sample of 6,649 firms and 19,332 loans in Column Specification (1) and for the opposite sample of 6,028 firms that receive 14,891 loans from either Relationship or Mixed Banks only in Column Specification (2). The effect is smaller in magnitude (2.0 and 1.6 percentage points for the full and opposite sample) but remains highly statistically significant, suggesting that the main result is not restricted to the sub-sample.

2.4.4.1 Additional Robustness Tests

Apart from alternative explanations, I also explore a battery of additional robustness tests in Table 2.10.[63] Panel B of Table 2.10 reports results for alternative non-performance measures, namely "Non-Performance" that equals one if a loan has any overdue payments or overdue days

[63]In unreported results, I confirm that robustness tests also hold when I use logit models instead or delinquencies over 90 days as the dependent variable.

and zero otherwise as well as "Non-Performance 180" that equals one if loans becomes delinquent for more than 180 days and zero otherwise. While there is no systematically significant effect of relationship lending on loan performance with firm characteristics, I find statistically significant effects with firm fixed effects that are similar in magnitude relative to their average values. The same firm is by around 40% and 56% more likely to become delinquent on any payments and for more than 180 days, respectively.

In Panel C of Table 2.10, I test for different loan characteristics. When splitting loans into local currency and US dollars loans in Rows "Local currency loans" and "US dollar loans" the effect becomes stronger for US dollars and loses significance and magnitude for local currency loans. This suggests that either Relationship Banks offer more flexibility in US dollars based on their dollar-denominated funding structure or that firms have more problems repaying US dollar denominated loans, having less income in US dollars. When splitting loans issued before and after 2011 in Rows "Loans between January 2009-2011" and "Loans after January 2011", the negative effect of SME loans of Mixed Banks mostly survives for loans after 2011, but is somewhat reduced for loans between January 2009-2011. This might be due to the reduced number of observations or generally higher non-performance for all banks during the crisis period.

In Row "W/o loans 50%" of Panel C, I exclude loans outside a range of 50% around the loan amount threshold that determines SME and corporate loans. I assume that, for Mixed Banks, the change from transaction-based to relationship lending does not switch from one to the other exactly at the threshold but happens gradually with larger loan amounts. This implies that differences in lending techniques between Relationship and Mixed Banks should be stronger further away from the threshold. As expected, the main effect continues to hold and is somewhat stronger in magnitude. For the last two Rows "Timing w/o SME & corp. Loans" and "Timing w/o SME Loans" , imagine a firm that receives a transaction-based loan, and then a relationship-based loan within the same bank (i.e, SME vs. corporate loan of the same Mixed Bank*)*. If the information sharing across departments does not work perfectly and the customer will be assigned, for example, to different loan officers in different departments, the timing of loans will not influence the main results. If information on customers is shared among

departments, then once a firm receives a transaction-based loan, the customer will be known to the bank and there will be no difference in loan performance once the firm receives a relationship-based loan. Therefore, in Rows "Timing w/o SME & corp. Loans" and "Timing w/o SME Loans" of Table 2.10, I exclude (i) all SME and corporate loans once a firm switches from SME to corporate loans and (ii) only SME loans once a firms switches from SME to corporate loans in the case of Mixed Banks. Only few loans are dropped and results remain virtually the same, suggesting imperfect information sharing across departments of Mixed Banks.

Panel D of Table 2.10 investigates different firm characteristics. In particular, I examine whether Relationship Banks are particularly helpful for firms in more opaque industries. Therefore, I distinguish between the main industries in Armenia such as trade, manufacturing and construction as well as other fields of services and other industries. While the negative effect of transaction-based lending on delinquencies disappears for more transparent industries in Rows "Trade, manufacturing, construction" and "Other industries", the effect increases to 3.3 and 3.6 percentage points in Row "Other fields of services", a more opaque industry. in Rows "New customers" and "Old customers", I distinguish between loans of new customers to a bank and customers with which a bank has had a previous relationship since Relationship Banks are more likely to know these old customers better and thus be more lenient towards them. In line with this prediction, the main effect increases to 2.7 percentage points for old customers but becomes almost completely insignificant for new customers.

The last Panel E of Table 2.10 compares different estimation techniques. In Row "Logit model", I estimate a logit model instead of a linear probability model for the benchmark specifications and report odds ratios to ease the interpretation. The odds of becoming delinquent on a relationship-based loans rise by 43% when controlling for firm characteristics and by 56% within the same firm relative to transaction-based lending. In Row "Matching on firm and loan amount", I rely on a matching technique similar to Ioannidou and Ongena (2010) and Beck et al. (2014b). Since matching is a nonparametric estimation technique, it does not impose a functional form on the relationship between matching variables and the dependent variable. In particular, I match SME loans of Relationship and Mixed Bank*s* of the same firm from the sub-sample of 4,441 loans and 10,656 loan-time observations. Then, I match on differences in

loan amounts between SME loans of Relationship and Mixed Banks, conditioning them to be in 0.5 or 0.25 standard deviations of the average loan amount of SME loans of the two bank types. In Row "Matching on firm and loan char.", I additionally match on differences in interest rates, maturity and credit ratings in the same way and exact matching of collateralized and non-collateralized loans. Specifications (1) and (2) in Table 2.10 correspond to a 0.5 and 0.25 standard deviation radius for each of the matching variables. While matching on firm and loan amounts yields similar results as for the main regression with an 2.6 percentage points effect, matching on more variables increases the negative effect to 4.3 and 6.2 percentage points, yet only significant in the former case which might be due to the small number of observations in the latter case. In the last two rows of Table 2.10, I implement clusters at bank level and double clustering at bank×time and firm level to account for the fact that observations within a bank and time period might be correlated. Although standard errors increase, the main effect remains significant.

2.5 Relationship Lending and Information Use

Since the measure of relationship lending is new to the literature and relies on survey responses, this section shows that it captures the use of soft information when pricing loans. Similar to Rajan, Seru, and Vig (2014) and Skrastins and Vig (2014), I assume that in a state with just hard information and no soft information available, hard information variables will perfectly predict the loan interest rate. In a state with additional soft information, hard information variables will not be able to completely explain interest rates and the unexplained part becomes a measure of soft information.

For the analysis, I use a regression model with multiplicative heteroskcedasticity introduced by Harvey (1976) and applied to banking by Cerqueiro, Degryse, and Ongena (2011). The model estimates mean effects on the interest rate and the determinants of the residual variance in interest rates. The model consists of an equation for the mean of interest rates, and a second one for the residual variance of interest rates:[64]

[64] A more detailed description of the methodology can be found in the Appendix C.

$$\text{Loan Spread}_{ijk} = \theta\text{'Controls}_{ijk} + \alpha_j + \varepsilon_{ijk}, \tag{2.5.1}$$

$$\begin{aligned} Log(\sigma^2_{ijk}) = \alpha_0 &+ \delta_1 \text{Corporate Loan}_{ijk} \\ &+ \delta_2 Mixed\,Bank_k \\ &+ \delta_3 \text{Corporate Loan}_{ijk} \times \text{Mixed Bank}_k, \end{aligned} \tag{2.5.2}$$

where i, j, k index loans, firms, and banks. Note that different from equation (2.3.1) I only use information at loan initiation such that each loan appears only once in the data set. The Loan Spread_{ijk} equals the loan interest rate minus the refinancing rate of the Armenian banks with the Armenian Central Bank. $Log(\sigma^2_{ijk})$ stands for the natural logarithm of the residual variance of the loan spread. The other variables are defined as in equation (2.3.1). By including loan contract terms as well as firm fixed effects, I control for all hard information variables that explain the variation in interest rates for the *same firm*. The remaining unexplained variation should capture the use of soft information. A positive effect on the variance of the unexplained part means that the variance increases, hard information variables are less predictive of interest rates and more soft information is used and vice versa for a negative effect.

The coefficient δ_2 estimates the effect of transactional lending on the variation in interest rates (soft information use) relative to relationship lending (SME loans of Mixed vs. Relationship Loans). The difference in coefficients δ_1 and δ_3 measures the effect on the variation in interest rates (soft information use) when both banks rely on relationship lending (corporate loans of Mixed vs. Relationship Loans).

Since I am interested in the effect of Relationship lending on the variation in interest rates, Table 2.11 only shows estimation results of the variance equation 2.5.2, where the columns correspond to the specifications of the columns in Table 2.3. Results on mean equation 2.5.1 are available upon request. All specifications in 2.11 reveal a negative effect of transactional lending on the variation in interest rates (SME loans of Mixed Banks). Transactional lending reduces the unexplained part of interest rates and leads to less use of soft information relative to relationship lending. A test for differences in coefficients shows that the use of soft information

does not significantly differ when both banks rely on relationship lending. In sum, these results suggest that the relationship lending measure indeed captures the use of soft information in loan pricing.

2.6 Conclusions

Although the empirical literature on relationship lending is quite extensive, few is known about the behavior of banks when firms are in distress. Combining survey data on banks' lending policies with unique credit registry data, this paper fills this gap by examining the effect of relationship lending on ex-post loan performance. In line with Von Thadden (1995) and Rajan (1992), I find that Relationship Banks tolerate temporary delinquencies without facing higher defaults and earning higher rents in the long run. When firms are in distress, Relationship Banks adjust contract terms and offer drawdowns on credit lines and overdrafts but do not rollover loans. Moreover, Relationship Banks are more likely to continue to lend to firms after past non-performance. The paper presents a new channel of how relationship lending serves as a liquidity insurance for firms in distress, offering greater financial flexibility and better access to finance.

The findings of the paper have several broader implications. Relying on soft information, relationship lending constitutes a critical tool to target SMEs which are the backbone of most economics.[65] This paper shows that relationship lending is especially beneficial when firms experience liquidity shortages. In the long run, firms will thus have longer investment horizons which should lead to more investments, employment and economic growth. Moreover, proprietary information seems to be key for banks to ease lending standards which calls for an improved contractual and informational framework. From a financial stability perspective, relationship lending appears to be an efficient lending technique to help firms in need without incurring higher losses for banks. Finally, the results might also be useful for other markets such as the labor market or insurance market in which close relationships help reduce existing information asymmetries.

[65] According to the website of the Global Alliance of SMEs, SMEs have provided nearly 50% of jobs in most countries (53% in US and 78% in Germany) and account for 75% and 39% of GDP in Germany and US (Global Alliance of SMEs (2014)).

References

Agarwal, Sumit, and Robert Hauswald, 2010, Distance and private information in lending, *Review of Financial Studies* 23, 2757–2788.

Allen, Franklin, 1990, The market for information and the origin of financial intermediation, *Journal of Financial Intermediation* 1, 3–30.

Baele, Lieven, Moazzam Farooq, and Steven Ongena, 2014, Of religion and redemption: Evidence from default on Islamic loans, *Journal of Banking & Finance* 44, 141–159.

Beck, Thorsten, Hans Degryse, Ralph De Haas, and Neeltje van Horen, 2014a, When arm's length is too far. Relationship banking over the business cycle, EBRD Working Paper No. 169.

Beck, Thorsten, Vasso Ioannidou, and Larissa Schaefer, 2014b, Foreigners vs. Natives: Bank lending technologies and loan pricing, Working Paper Series Cass Business School No. 28.

Berg, Tobias, 2014, Playing the devil's advocate: The causal effect of risk management on loan quality, Working Paper,Available at SSRN: http://ssrn.com/abstract=2435158.

Berger, Allen N., and Gregory F. Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of Business* 68, 351–381.

Berlin, Mitchell, and Loretta Jean Mester, 1999, Deposits and relationship lending, *Review of Financial Studies* 12, 579–607.

Bolton, Patrick, Xavier Freixas, Leonardo Gambacorta, and Paolo Emilio Mistrulli, 2013, Relationship and transaction lending in a crisis, NBER Working Paper 19467.

Boot, Arnoud W.A., 2000, Relationship banking: What do we know?, *Journal of Financial Intermediation* 9, 7–25.

Boot, Arnoud W.A., and Anjan V. Thakor, 1994, Moral hazard and secured lending in an infinitely repeated credit market game, *International Economic Review* 35, 899–920.

Cameron, A. Colin, and Pravin K. Trivedi, 2005, *Microeconometrics: methods and applications* (Cambridge University Press).

CBA, 2014, Financial Stability Report 2013, Technical report, Central Bank of The Republic of Armenia.

Cerqueiro, Geraldo, Hans Degryse, and Steven Ongena, 2011, Rules versus discretion in loan rate setting, *Journal of Financial Intermediation* 20, 503–529.

De la Torre, Augusto, María Soledad Martínez Pería, and Sergio L Schmukler, 2010, Bank involvement with SMEs: Beyond relationship lending, *Journal of Banking & Finance* 34, 2280–2293.

Degryse, Hans, Moshe Kim, and Steven Ongena, 2009, *Microeconometrics of banking: Methods, applications, and results* (Oxford University Press).

Degryse, Hans, and Steven Ongena, 2005, Distance, lending relationships, and competition, *Journal of Finance* 60, 231–266.

Diamond, Douglas W., 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 51, 393–414.

Elsas, Ralf, and Jan Pieter Krahnen, 1998, Is relationship lending special? Evidence from credit-file data in Germany, *Journal of Banking & Finance* 22, 1283–1316.

Fahlenbrach, Rüdiger, Robert Prilmeier, and Rene M. Stulz, 2012, This time is the same: Using bank performance in 1998 to explain bank performance during the recent financial crisis, *Journal of Finance* 67, 2139–2185.

Garmaise, Mark J., and Gabriel Natividad, 2014, Negative financial spillovers in local banking negative financial spillovers in local banking markets, Working Paper New York University.

Global Alliance of SMEs, GASME, 2014, Four global experience to support SME development, available at http://www.globalsmes.org.

Harhoff, Dietmar, and Timm Körting, 1998, Lending relationships in Germany – Empirical evidence from survey data, *Journal of Banking & Finance* 22, 1317–1353.

Harvey, Andrew C, 1976, Estimating regression models with multiplicative heteroscedasticity, *Econometrica: Journal of the Econometric Society* 44, 461–465.

Haselmann, Rainer, David Schoenherr, and Vikrant Vig, 2013, Lending in social networks, Working Paper London Business School.

IFC, 2013, Verifying accuracy of IFC's SME measurement, World Bank and International Finance Corporation.

Ioannidou, Vasso, and Steven Ongena, 2010, Time for a change: Loan conditions and bank behavior when firms switch banks, *Journal of Finance* 65, 1847–1877.

Keys, Benjamin J, Tanmoy K Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? Evidence from subprime loans, *Quarterly Journal of Economics* 125, 307–362.

Kysucky, Vlado, and Lars Norden, 2014, The benefits of relationship lending in a cross-country context: A meta-analysis, *Management Science, forthcoming* .

Li, Yan, Ruichang Lu, and Anand Srinivasan, 2013, Relationship bank behavior during borrower distress, Working Paper, Available at SSRN: http://ssrn.com/abstract=1773069.

Liberti, Jose M, and Atif R Mian, 2009, Estimating the effect of hierarchies on information use, *Review of Financial Studies* 22, 4057–4090.

Machauer, Achim, and Martin Weber, 1998, Bank behavior based on internal credit ratings of borrowers, *Journal of Banking & Finance* 22, 1355–1383.

McCrary, Justin, 2008, Manipulation of the running variable in the regression discontinuity design: A density test, *Journal of Econometrics* 142, 698–714.

Morales, Paola, 2014, Strategic choice of delinquencies under firm liquidity constraints, Working Paper, Tilburg University.

Petersen, Mitchell A., and Raghuram G. Rajan, 1994, The benefits of lending relationships: Evidence from small business data, *Journal of Finance* 49, 3–37.

Petersen, Mitchell A., and Raghuram G. Rajan, 1995, The effect of credit market competition on lending relationships, *Quarterly Journal of Economics* 407–443.

Puri, Manju, Jörg Rocholl, and Sascha Steffen, 2013, What kinds of bank-client relationships matter in reducing loan defaults and why?, Working Paper, Available at SSRN: http://ssrn.com/abstract=1572673.

Rajan, Raghuram G., 1992, Insiders and outsiders: The choice between informed and arm's-length debt, *Journal of Finance* 47, 1367–1400.

Rajan, Uday, Amit Seru, and Vikrant Vig, 2014, The failure of models that predict failure: Distance, incentives and defaults, *Journal of Financial Economics, forthcoming* .

Ramakrishnan, Ram T.S., and Anjan V. Thakor, 1984, Information reliability and a theory of financial intermediation, *Review of Economic Studies* 51, 415–432.

Schoar, Antoinette, 2014, The personal side of relationship banking, Working Paper MIT.

Sharpe, Steven A., 1990, Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships, *Journal of Finance* 45, 1069–1087.

Skrastins, Janis, and Vikrant Vig, 2014, How organizational hierarchy affects information production, Working Paper London Business School.

SME DNC, 2010, Reference Book: Small and Medium Entrepreneurship Sector in Armenia 2007-2009, Small and Medium Entrepreneurship Development National Center of Armenia.

Stein, Jeremy C, 2002, Information production and capital allocation: Decentralized versus hierarchical firms, *The Journal of Finance* 57, 1891–1921.

Von Thadden, Ernst-Ludwig, 1995, Long-term contracts, short-term investment and monitoring, *Review of Economic Studies* 62, 557–575.

World Bank, 2011, Armenia: Country Profile, Enterprise Survey: Running a business in Armenia, World Bank and International Finance Corporation, Rev. 7.

World Bank, 2014a, Country Program Snapshot, World Bank and IFC.

Tables

Table 2.1: Bank Level Descriptive Statistics

The Table reports bank level summary statistics for 15 Armenian banks on the importance of relationship lending and the average number of hierarchical layers for loan approval by loan type, the average loan amount threshold in US$, the market share in terms of loan number and loan value, the average borrower size based on the total borrowing amount across all banks in US$ between January 2009 and June 2013 as well as bank size based on total assets as of 2009 and foreign ownership that equals one if more that 50% of equity is foreign-owned.

		SME Loans		Corporate Loans							
Bank ID	Bank Type	Relationship Lending	Average Hierarchy	Relationship Lending	Average Hierarchy	Average Threshold in $US	Share in Number of Loans (%)	Share in Value of Loans (%)	Borrower Size	Bank Size	Foreign
38	Mixed	3	2	5	2	2,751,486	0.027	0.029	1,727,395	Small	0
59	Mixed	4	2	5	2	1,500,000	0.065	0.216	3,033,712	Big	0
70	Relationship	5	1	5	1	518,444	0.061	0.041	1,0679,69	Medium	0
219	Mixed	4	3	5	1	166,639	0.076	0.032	529,078	Medium	0
274	Relationship	5	2	5	3	837,301	0.054	0.036	466,309	Medium	1
457	Mixed	4	2	5	3	2,751,486	0.089	0.136	2,133,701	Big	0
470	Mixed	4	3	5	3	500,000	0.169	0.067	721,984	Medium	0
520	Relationship	5	2	5	1	1,000,000	0.037	0.108	3,756,655	Big	0
523	Relationship	5	2	5	2	1,355,53	0.025	0.017	1,211,437	Medium	0
662	Relationship	5	1	5	2	500,000	0.137	0.068	313,301	Big	0
702	Relationship	5	1	5	1	271,106	0.122	0.055	790,899	Medium	0
772	Mixed	4	1	5	3	661,568	0.026	0.024	1,833,290	Small	0
776	Relationship	5	3	5	3	300,000	0.040	0.028	1,748,996	Small	1
798	Relationship	5	2	5	2	200,000	0.037	0.029	1,319,247	Small	0
995	Relationship	5	5	5	5	500,000	0.035	0.112	2,471,792	Big	1

Table 2.2: Loan Level Descriptive Statistics

The Table reports loan level summary statistics on ex-post loan performance measures, loan characteristics, relationship characteristics, and firm characteristics by bank type between January 2009 and June 2013. The two bank types are *Relationship Banks* that always rely on relationship lending and *Mixed Banks* that rely on relationship lending only for corporate loans. Definitions of the variables can be found in Table A.1 of Appendix A. The left panel "Total Sample" reports summary statistics for the total sample of 19,332 loans to 6,649 firms. The right panel "Sub-Sample" reports summary statistics for a sub-sample of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks*. The Columns "Difference *t*-test" in both panels report *t*-statistics for differences in means between the two bank types and indicate significance at the 1%, 5%, and 10% levels with ***, **, *.

	Total Sample					Sub-Sample				
	Relationship Bank		Mixed Bank		Difference	Relationship Bank		Mixed Bank		Difference
Variable Names	Mean	Std	Mean	Std	*t*-test	Mean	Std	Mean	Std	*t*-test
Ex Post Loan Performance										
Non-Performance	0.058	0.233	0.061	0.238	-0.002	0.051	0.220	0.046	0.209	0.005
Non-Performance 0-90 days	0.100	0.300	0.075	0.263	0.026***	0.088	0.284	0.058	0.234	0.030***
Non-Performance 90 days	0.046	0.210	0.037	0.188	0.010***	0.041	0.199	0.023	0.150	0.018***
Non-Performance 180 days	0.029	0.168	0.021	0.144	0.008***	0.025	0.156	0.015	0.123	0.010**
Default (loss/written-off)	0.020	0.139	0.012	0.110	0.008***	0.012	0.109	0.009	0.095	0.003
Loan Characteristics										
Credit Classification	4.99	0.17	4.99	0.14	-0.002	4.98	0.18	4.99	0.13	-0.007
Interest Rate	15.37	3.88	15.10	3.72	0.264***	13.81	3.73	13.76	3.56	0.045
Loan Spread	8.00	4.37	7.62	4.08	0.380***	6.41	4.21	6.25	3.98	0.152
Loan Amount in US$	181,386	606,152	224,708	714,815	-43,322***	352,903	963,534	341,802	752,242	11,100
Collateral	0.82	0.38	0.88	0.32	-0.062***	0.76	0.43	0.87	0.33	-0.108***
Guarantee	0.05	0.22	0.11	0.31	-0.053***	0.11	0.31	0.16	0.37	-0.051***
Loan Maturity in Months	36.32	21.37	33.38	22.46	2.940***	32.39	22.65	32.00	24.42	0.389
Other Loan Characteristics										
Loan Location in Yerevan	0.59	0.49	0.75	0.43	-0.163***	0.72	0.45	0.77	0.42	-0.051***
Wholesale and Retail Trade Loan	0.46	0.50	0.42	0.49	0.041***	0.51	0.50	0.44	0.50	0.070***
Other Fields of Service Loan	0.13	0.34	0.22	0.41	-0.083***	0.11	0.31	0.14	0.35	-0.037***
Loan in USD	0.46	0.50	0.51	0.50	-0.051***	0.58	0.49	0.61	0.49	-0.026*

Table 2.2 (continued): Loan Level Descriptive Statistics

	Total Sample					Sub-Sample				
	Relationship Bank		Mixed Bank		Difference	Relationship Bank		Mixed Bank		Difference
Variable Names	Mean	Std	Mean	Std	*t*-test	Mean	Std	Mean	Std	*t*-test
Relationship Characteristics										
Relationship in Months	15.63	18.16	15.92	17.65	-0.291	17.43	19.27	16.22	17.27	1.211**
Scope	0.18	0.38	0.24	0.43	-0.062***	0.30	0.46	0.31	0.46	-0.012
Primary Bank	0.93	0.26	0.90	0.30	0.032***	0.79	0.41	0.74	0.44	0.049***
Number of Relationships	1.93	1.39	1.94	1.47	-0.012	3.41	2.08	3.24	2.01	0.176***
Multiple Relationships	0.53	0.50	0.51	0.50	0.025***	1	0	1	0	0
Firm Characteristics										
Firm Location in Yerevan	0.59	0.49	0.75	0.43	-0.163***	0.72	0.45	0.77	0.42	-0.051***
Wholesale Retail Trade	0.23	0.42	0.16	0.36	0.075***	0.24	0.43	0.18	0.39	0.057***
Industry Firm										
Other Fields of Service	0.54	0.50	0.63	0.48	-0.087***	0.54	0.50	0.60	0.49	-0.057***
Industry Firm										
Private Firm	0.54	0.50	0.65	0.48	-0.104***	0.53	0.50	0.61	0.49	-0.076***
Observations	10,598		8,734		19,332	2,151		2,290		4,441

Table 2.3: Relationship Lending and Loan Performance (NPL 0-90 days)

The Table reports regression results from a linear probability model for a sub-sample of 10,656 loan-time observations of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. The dependent variable is Loan Performance$_{ijkt}$ that equals one when a loan is delinquent for less than 90 days. The main independent variable is "Mixed Bank" which measures the performance of transaction-based relative to relationship-based loans, i.e., SME loans of *Mixed* versus *Relationship Banks* (the reference group). Columns (1) reports results for with loan characteristics and firm characteristics. In Columns (2)-(5) firm fixed effects are added. Column (3) adds loan origination fixed effects, Column (4) bank characteristics, and Column (5) other loan characteristics. Column (6) introduces firm×time fixed effect and thus the sample is reduced to 3,790 loan-time observations of 1,952 loans to 318 firms. The last row presents p-values for a Wald test of differences in coefficients when both banks rely on relationship lending, i.e., corporate loans of *Relationship* versus *Mixed Banks*. Definitions of the variables can be found in Table A.1 of Appendix A. Standard errors are clustered at firm level and presented in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10%.

	(1)	(2)	(3)	(4)	(5)	(6)
Bank Lending Technologies						
Corporate Loan	0.035*	0.025	0.025	0.020	0.025	0.021
(Relationship Lending)	(0.020)	(0.017)	(0.016)	(0.016)	(0.016)	(0.020)
Mixed Bank	-0.025***	-0.025***	-0.025***	-0.032***	-0.024***	-0.027**
(Transactional Lending)	(0.009)	(0.008)	(0.008)	(0.012)	(0.009)	(0.013)
Corporate Loan × Mixed Bank	-0.021	0.004	0.001	0.009	0.004	-0.001
(Relationship Lending)	(0.021)	(0.017)	(0.017)	(0.017)	(0.017)	(0.020)
Loan Characteristics						
Credit Rating	-0.191***	-0.178***	-0.178***	-0.175***	-0.182***	-0.091
	(0.061)	(0.062)	(0.062)	(0.062)	(0.062)	(0.154)
Credit Interest Rate	0.005***	0.002	0.003*	0.001	0.003	0.008*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)
Loan Amount	0.004	-0.002	-0.001	-0.003	-0.003	0.008
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)
Collateral	0.029***	0.015	0.012	0.015	0.018*	-0.001
	(0.009)	(0.010)	(0.012)	(0.011)	(0.010)	(0.023)
Guarantee	0.006	0.001	-0.001	-0.001	0.003	-0.036
	(0.017)	(0.025)	(0.025)	(0.026)	(0.025)	(0.023)
Maturity	0.016***	0.013**	0.015**	0.013**	0.012**	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)

Table 2.3 (continued): Relationship Lending and Loan Performance (NPL 0-90 days)

	(1)	(2)	(3)	(4)	(5)	(6)
Firm Characteristics						
Firm Location Yerevan	0.004					
	(0.010)					
Wholesale Retail Trade Firm	-0.011					
	(0.014)					
Other Fields of Service Firm	0.011					
	(0.013)					
Private Firm	-0.001					
	(0.010)					
Bank Characteristics						
Large Bank				0.012		
				(0.012)		
Foreign Bank				-0.007		
				(0.021)		
High Average Hierarchy				0.021*		
				(0.012)		
Other Loan Characteristics						
Loan Location in Yerevan					0.007	
					(0.019)	
Wholesale Retail Trade Loan					-0.002	
					(0.013)	
Other Fields of Service Loan					-0.023*	
					(0.013)	
Loan in USD					0.015	
					(0.009)	
Fixed Effects						
Firm Fixed Effects	No	Yes	Yes	Yes	Yes	No
Loan Origination Fixed Effects	No	No	Yes	No	No	No
Firm×Time Fixed Effects	No	No	No	No	No	Yes
Constant	0.811**	0.891***	0.838***	0.876***	0.889***	0.319
	(0.315)	(0.318)	(0.311)	(0.318)	(0.321)	(0.796)
R-squared	0.044	0.293	0.294	0.293	0.294	0.633
Observations (Loan-Time Level)	10,656	10,656	10,656	10,656	10,656	3,790
Corporate Loan = Corporate Loan × Mixed Bank	0.144	0.498	0.444	0.711	0.484	0.525

Table 2.4: Relationship Lending and Loan Performance (NPL 90 days)

The Table reports regression results from a linear probability model for a sub-sample of 4,441 loans to 271 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. The dependent variable is Loan Performance$_{ijkt}$ that equals one when a loan is delinquent for more than 90 days. The main independent variable is "Mixed Bank" which measures the performance of transaction-based relative to relationship-based loans, i.e., SME loans of *Mixed* versus *Relationship Banks* (the reference group). Columns (1) reports results for with loan characteristics and firm characteristics. In Columns (2)-(5) firm fixed effects are added. Column (3) adds loan origination fixed effects, Column (4) bank characteristics, and Column (5) other loan characteristics. Column (6) introduces firm×time fixed effect and thus the sample is reduced to 3,790 loan-time observations of 1,952 loans to 318 firms. The last row presents p-values for a Wald test of differences in coefficients when both banks rely on relationship lending, i.e., corporate loans of *Relationship* versus *Mixed Banks*. Definitions of the variables can be found in Table A.1 of Appendix A.. Standard errors are clustered at firm level and presented in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10%.

	(1)	(2)	(3)	(4)	(5)	(6)
Bank Lending Technologies						
Corporate Loan	0.004	0.001	-0.000	-0.002	0.002	-0.001
(Relationship Lending)	(0.017)	(0.008)	(0.008)	(0.009)	(0.008)	(0.010)
Mixed Bank	-0.016***	-0.016***	-0.016***	-0.019***	-0.015***	-0.012
(Transactional Lending)	(0.006)	(0.005)	(0.005)	(0.007)	(0.005)	(0.007)
Corporate Loan × Mixed Bank	0.012	0.021*	0.021*	0.022*	0.020*	-0.000
(Relationship Lending)	(0.023)	(0.011)	(0.011)	(0.012)	(0.011)	(0.013)
Loan Characteristics						
Credit Rating	-0.388***	-0.206***	-0.208***	-0.206***	-0.209***	-0.146***
	(0.037)	(0.044)	(0.044)	(0.044)	(0.045)	(0.039)
Credit Interest Rate	0.004***	-0.000	0.001	0.000	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Loan Amount	0.007**	-0.003*	-0.003*	-0.003	-0.003**	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Collateral	0.021**	0.019**	0.018***	0.020***	0.020***	0.016
	(0.009)	(0.007)	(0.006)	(0.007)	(0.008)	(0.013)
Guarantee	-0.002	0.003	0.003	0.004	0.004	-0.007
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.011)
Maturity	0.002	0.005*	0.005*	0.006**	0.005	0.003
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)

Table 2.4 (continued): Relationship Lending and Loan Performance (NPL 90 days)

	(1)	(2)	(3)	(4)	(5)	(6)
Firm Characteristics						
Firm Location Yerevan	0.009					
	(0.008)					
Wholesale Retail Trade Firm	0.004					
	(0.014)					
Other Fields of Service Firm	-0.002					
	(0.012)					
Private Firm	0.010					
	(0.009)					
Bank Characteristics						
Large Bank				-0.005		
				(0.008)		
Foreign Bank				-0.003		
				(0.009)		
High Average Hierarchy				0.007		
				(0.007)		
Loan Location in Yerevan					0.002	
					(0.011)	
Other Loan Characteristics						
Wholesale Retail Trade Loan					-0.002	
					(0.009)	
Other Fields of Service Loan					-0.017*	
					(0.010)	
Loan in USD					0.007	
					(0.006)	
Fixed Effects						
Firm Fixed Effects	No	Yes	Yes	Yes	Yes	No
Loan Origination Fixed Effects	No	No	Yes	No	No	No
Firm×Time Fixed Effects	No	No	No	No	No	Yes
Constant	1.800***	1.069***	1.061***	1.057***	1.072***	0.720***
	(0.184)	(0.215)	(0.220)	(0.212)	(0.215)	(0.193)
R-squared	0.186	0.480	0.482	0.480	0.481	0.847
Observations (Loan-Time Level)	10,656	10,656	10,656	10,656	10,656	3,790
Corporate Loan = Corporate Loan × Mixed Bank	0.819	0.253	0.208	0.179	0.273	0.953

Table 2.5: Relationship Lending and Long-Term Performance

The Table reports performance statistics of SME loans selected from the total sample of 19,332 loans to 6,649 firms and the sub-sample of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. For SME loans *Relationship Banks* rely on relationship lending and *Mixed Banks* on transactional lending. Both panels show loan defaults (loss/written-off), recovery rates, the percentage of the loan and interest amount not repaid in time or lost/written-off for SME loans that are observed until maturity. Loan default equals one if a loan is in loss or written-off status at the end of the loan spell and zero otherwise. Recovery rate equals one if a loan has been delinquent during the loan spell but did not default at the end of the loan spell. The percentage of loan and interest amount not repaid in time (lost/written-off) stands for the ratio of the principal and interest rate amount over the total contract amount in case of delinquencies below/over 90 days or default, conditional on non-performance at the end of the loan spell. While Panel A conditions loans to have been delinquent for less than 90 days during the loan spell, Panel B conditions loans to have been delinquent over 90 days during the loan spell. For all panels, the column "Difference *t*-test" reports *t*-statistics for differences in means between the two bank types and indicates significance at the 1%, 5%, and 10% levels with ***, **, *.

Panel A: Defaults, Recovery Rates, and Losses for SME Loans if NPL 0-90 days

	Total Sample: SME Loans					Sub-Sample: SME Loans				
	Relationship Banks		Mixed Banks		Difference	Relationship Banks		Mixed Banks		Difference
Variable Names	Mean	Std	Mean	Std	*t*-test	Mean	Std	Mean	Std	*t*-test
Loan Default (loss/written-off)	0.249	0.433	0.198	0.399	0.051*	0.169	0.377	0.212	0.412	-0.043
Recovery Rate	0.127	0.333	0.158	0.366	-0.031	0.156	0.365	0.173	0.382	-0.017
% of loan and interest amount not repaid in time	0.047	0.127	0.104	0.229	-0.057***	0.064	0.185	0.111	0.240	-0.047
% of lost/written-off loan and interest amount	0.256	0.858	0.181	0.551	0.075	0.145	0.675	0.147	0.374	-0.002
Observations (Loan Level)	543		303		846	77		52		129

Panel B: Defaults, Recovery Rates, and Losses for SME Loans if NPL 90 days

	Total Sample: SME Loans					Sub-Sample: SME Loans				
	Relationship Banks		Mixed Banks		Difference	Relationship Banks		Mixed Banks		Difference
Variable Names	Mean	Std	Mean	Std	*t*-test	Mean	Std	Mean	Std	*t*-test
Loan Default (loss/in written-off status)	0.608	0.489	0.522	0.501	0.087*	0.512	0.506	0.600	0.498	-0.088
Recovery Rate	0.226	0.419	0.296	0.458	-0.070*	0.220	0.419	0.333	0.479	-0.114
% of loan and interest amount not repaid in time	0.849	1.407	0.931	1.137	-0.082	0.907	1.834	0.608	0.619	0.299
% of lost/written-off loan and interest amount	0.672	1.328	0.559	0.938	0.113	0.661	1.688	0.340	0.453	0.321
Observations (Loan Level)	301		186		487	41		30		71

Table 2.6: Relationship Lending and Long-Term Rents

The Table reports summary statistics and regression results of returns on loans for SME loans selected from the total sample of 19,332 loans to 6,649 firms and the sub-sample of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. For SME loans *Relationship Banks* rely on relationship lending and *Mixed Banks* on transactional lending. Returns on loans are defined in equation 2.4.1 as the value-weighted interest rate and loss of a bank in case of non-performance. The loss of a bank is defined as the overdue principal plus interest rate amount over the contract amount. The first panel shows the return on SME loans on loan level for the total sample and the sub-sample by bank type. The second and third panels show the return on SME loans that have been delinquent for less than 90 days and over 90 days. For all panels, the column "Difference *t*-test" reports *t*-statistics for differences in means between the two bank types and indicates significance at the 1%, 5%, and 10% levels with ***, **, *. The last panel shows results from a regressions of return on loans on a *Mixed Bank* dummy without firm fixed effects for all SME loans ("All Loans"), for SME loans not delinquent for less than 90 days ("NPL 0-90=0"), for SME loans delinquent for less than 90 days ("NPL 0-90=1"), for SME loans not delinquent over 90 days ("NPL 90=0"), and for SME loans delinquent over 90 days ("NPL 90=1").

	Total Sample: SME Loans					Sub-sample: SME Loans				
	Relationship Banks		Mixed Banks		Difference	Relationship Banks		Mixed Banks		Difference
Variable Names	Mean	Std	Mean	Std	*t*-test	Mean	Std	Mean	Std	*t*-test
Return on loans										
Return on Loans	15.08	4.514	14.7	4.46	0.385***	13.75	4.421	13.53	4.201	0.221
Observations (Loan-Level)	9,630		8,289		17,919	1,770		2,091		3,869
Return on loans given NPL 0-90 days										
Return on Loans	13.22	5.955	10.68	6.127	2.544***	11.95	5.495	9.521	6.048	2.443***
Observations (Loan-Level)	950		620		1,570	138		117		255
Return on loans given NPL 90 days										
Return on Loans	7.819	5.137	7.265	4.877	0.554	7.63	5.086	6.368	4.419	1.262
Observations (Loan-Level)	421		294		715	67		43		110

Regressions on return on loans

	Total Sample: SME Loans			Sub-sample: SME Loans		
	All Loans	NPL 0-90=1	NPL 90=1	All Loans	NPL 0-90=1	NPL 90=1
Mixed Bank	-0.385**	-2.544***	-0.554	-0.221	-2.443**	-1.262
	(0.159)	(0.505)	(0.501)	(0.356)	(1.209)	(1.179)
Constant	15.080***	13.220***	7.819***	13.746***	11.955***	7.630***
	(0.104)	(0.236)	(0.281)	(0.249)	(0.564)	(0.718)
R-squared	0.002	0.041	0.003	0.001	0.043	0.016
Observations (Loan-Level)	17,919	1,570	715	3,861	255	110

Table 2.7: Relationship Lending and Bank Behavior

The Table reports summary statistics loans selected from the total sample of 19,332 loans to 6,649 firms and the sub-sample of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. Panel A reports summary statistics on past non-performance of any kind with the present bank of the firm, with any bank including the former and with other banks of the firm collapsed to bank-firm-time and bank-firm level. Panel B reports summary statistics on rollover loans as well as increases and decreases in the interest rate, loan amount and maturity during the loan spell for the total sample and sub-sample of SME loans given delinquencies below 90 days. In particular, I declare a loan of a firm to be a rollover loan if the loan has been non-performing and a new loan has been issued within one, two, or three months before or after the non-performing loan with the same bank. Changes in contract terms are measured at the period and up to three periods after delinquencies below 90 days on loan-time level and then collapsed to loan level. Panel C repeats the same analysis given delinquencies over 90 days. Panel D reports whether firms that have credit lines or overdrafts with the same bank draw down on them in periods of delinquencies and up to three periods later on loan-time level and then collapsed to loan level for the total sample of SME loans. For SME loans *Relationship Banks* rely on relationship lending and *Mixed Banks* on transactional lending. For all panels, the column "Difference t-test" reports t-statistics for differences in means between the two bank types and indicates significance at the 1%, 5%, and 10% levels with ***, **, *.

Panel A: Distribution of Ex-ante Borrower Risks

	Total Sample: Bank-Firm-Time Level					Total Sample: Bank-Firm Level				
	Relationship Bank		Mixed Bank		Difference	Relationship Bank		Mixed Bank		Difference
Variable Names	Mean	Std	Mean	Std	t-test	Mean	Std	Mean	Std	t-test
Past NPL with Bank	0.016	0.126	0.011	0.103	0.005***	0.013	0.104	0.008	0.080	0.005**
Past NPL with Any Bank	0.017	0.128	0.018	0.132	-0.001	0.014	0.116	0.017	0.124	-0.002
Past NPL with Other Banks	0.015	0.119	0.016	0.124	-0.001	0.013	0.110	0.015	0.121	-0.002
Observations	7,874		6,322		14,196	4,531		3,352		7,883

Panel B: Rollover Loans and Renegotiations of Contract Terms given NPL 0-90 days

	Total Sample: SME Loans					Sub-Sample: SME Loans				
	Relationship Bank		Mixed Bank		Difference	Relationship Bank		Mixed Bank		Difference
Variable Names	Mean	Std	Mean	Std	t-test	Mean	Std	Mean	Std	t-test
Rollover loans (-\+ 1 months)	0.005	0.072	0.005	0.069	0.000	0.007	0.085	0.009	0.093	-0.001
Rollover loans (-\+ 2 months)	0.008	0.091	0.008	0.090	0.000	0.022	0.146	0.009	0.093	0.013
Rollover loans (-\+ 3 months)	0.012	0.107	0.015	0.120	-0.003	0.022	0.146	0.017	0.130	0.005
Increase in Interest Rate	0.053	0.223	0.032	0.177	0.020*	0.044	0.205	0.051	0.222	-0.008
Increase in Amount	0.004	0.065	0.008	0.090	-0.004	0.015	0.120	0.009	0.093	0.006
Increase in Maturity	0.041	0.199	0.015	0.120	0.0267***	0.044	0.205	0.000	0.000	0.0435**
Decrease in Interest Rate	0.053	0.223	0.027	0.163	0.025**	0.051	0.220	0.026	0.159	0.025
Decrease in Amount	0.005	0.072	0.008	0.090	-0.003	0.015	0.120	0.000	0.000	0.015
Decrease in Maturity	0.028	0.166	0.011	0.106	0.017**	0.022	0.146	0.000	0.000	0.022
Observations (Loan Level)	950		620		1,570	138		117		255

Table 2.7: Relationship Lending and Bank Behavior

Panel C: Rollover Loans and Renegotiations of Contract Terms given NPL 0-90 days

	Total Sample: SME Loans					Sub-Sample: SME Loans				
	Relationship Bank		Mixed Bank		Difference	Relationship Bank		Mixed Bank		Difference
Variable Names	Mean	Std	Mean	Std	*t*-test	Mean	Std	Mean	Std	*t*-test
Rollover loans (-\+ 1 months)	0.005	0.069	0.003	0.058	0.001	0.000	0.000	0.000	0.000	0.000
Rollover loans (-\+ 2 months)	0.012	0.108	0.003	0.058	0.008	0.030	0.171	0.000	0.000	0.030
Rollover loans (-\+ 3 months)	0.012	0.108	0.014	0.116	-0.002	0.030	0.171	0.047	0.213	-0.017
Increase in Interest Rate	0.097	0.297	0.024	0.153	0.074***	0.104	0.308	0.047	0.213	0.058
Increase in Amount	0.007	0.084	0.007	0.082	0.000	0.015	0.122	0.000	0.000	0.015
Increase in Maturity	0.052	0.223	0.020	0.142	0.032**	0.060	0.239	0.000	0.000	0.060
Decrease in Interest Rate	0.105	0.306	0.027	0.163	0.077***	0.104	0.308	0.023	0.152	0.081
Decrease in Amount	0.010	0.097	0.010	0.101	-0.000	0.030	0.171	0.000	0.000	0.030
Decrease in Maturity	0.045	0.208	0.007	0.082	0.038***	0.045	0.208	0.000	0.000	0.045
Observations (Loan Level)	421		294		715	67		43		110

Panel D: Drawdown on Credit Lines and Overdrafts given Non-Performance

	Total Sample: SME loans				
	Relationship Bank		Mixed Bank		Difference
	Mean	Std	Mean	Std	*t*-test
Drawdown on Credit Lines/Overdrafts given NPL 0-90 days					
Drawdown on Credit Lines/Overdrafts	0.398	0.491	0.278	0.451	0.119*
Observations (Loan Level)	166		79		245
Drawdown on Credit Lines/Overdrafts given NPL 90 days					
Drawdown on Credit Lines/Overdrafts	0.368	0.486	0.231	0.43	0.137
Observations (Loan Level)	68		26		94
Drawdown on Credit Lines/Overdrafts given NPL 0-90 or 90 day					
Drawdown on Credit Lines/Overdrafts	0.422	0.495	0.315	0.467	0.108*
Observations (Loan Level)	187		89		276
Drawdown on Credit Lines/Overdrafts given NPL					
Drawdown on Credit Lines/Overdrafts	0.18	0.384	0.076	0.265	0.104***
Observations (Loan Level)	384		211		595

Table 2.8: Use of Credit Registry for Loan Monitoring and Granting Purposes

The Table reports regression results for a sample of 2,737 firms that exclusively received loans either from *Relationship* or *Mixed Banks* and for which data on banks' use of the credit registry for loan monitoring and granting purposes exists from June 2012 until June 2013. The dependent variables are either "Monitoring" or "Granting", indicating the number of bank inquiries of a firm's credit registry information for loan monitoring or granting purposes. The main independent variable is "Mixed Bank Firm" which measures whether a firm that received loans only from *Mixed Banks* was inquired through the credit registry for loan monitoring or granting purposes. The reference group are firm that received loans only from *Relationship Banks* and were inquired by these. To control for differences in firm characteristics, I add firm size based on the average total outstanding debt across banks and past non-performance measures with their bank, any bank and all banks in columns (2) and (4) as well as firm location, industry and ownership fixed effects in columns (3) and (4). Standard errors are robust and presented in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10%.

	# of Loan Monitoring Inquiries			# of Loan Granting Inquiries		
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed Bank Firm	-0.334***	-0.257**	-0.263**	-0.037	0.006	0.004
	(0.109)	(0.109)	(0.119)	(0.062)	(0.061)	(0.065)
Firm Size		0.303***	0.319***		0.146***	0.154***
		(0.034)	(0.037)		(0.018)	(0.019)
Past NPL with Bank		-0.182	-0.229		-0.143	-0.153
		(0.259)	(0.286)		(0.186)	(0.196)
Past NPL with Any Bank		0.207	0.280		0.274	0.260
		(0.773)	(0.816)		(0.602)	(0.619)
Past NPL with Other Banks		1.176	1.033		0.210	0.272
		(1.112)	(1.198)		(0.654)	(0.658)
Fixed Effects						
Firm Location Fixed Effects	No	No	Yes	No	No	Yes
Firm Industry Fixed Effects	No	No	Yes	No	No	Yes
Firm Ownership Fixed Effects	No	No	Yes	No	No	Yes
Constant	2.205***	-1.340***	-2.097***	2.042***	0.439**	0.341
	(0.061)	(0.386)	(0.797)	(0.038)	(0.194)	(0.540)
R-squared	0.008	0.122	0.160	0.000	0.041	0.078
Observations	1,295	1,295	1,295	2,245	2,245	2,245

Table 2.9: Overdue Loan Repayments

The Table reports regression results for a sub-sample of 1,193 loan-time observations of 323 loans to 126 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013 and had overdue principal and interest rate repayments for less than 90 days. In the first two columns the dependent variable is the overdue principal and interest rate amount over the contract amount, the following two columns use the same numerator but set it relative to the total outstanding debt of the firm with the respective bank in a period and the last two columns use the total outstanding debt of a firm in a period. The main independent variable is "SME Loan×Mixed Bank" which measures the overdue exposure of transaction-based relative to relationship-based loans, i.e., SME loans of *Mixed* versus *Relationship Banks* (the reference group). For each dependent variable, I use firm characteristics or firm fixed effects to control for firm heterogeneity. Standard errors are robust and presented in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10%.

	NPL Amount over Contract Amount		NPL Amount over Bank Exposure		NPL Amount over Total Exposure	
	(1)	(2)	(3)	(4)	(5)	(6)
Corporate Loan	-0.013	-0.007	-0.204	-0.647	-0.012*	-0.003
(Relationship Lending)	(0.016)	(0.015)	(0.197)	(0.685)	(0.007)	(0.009)
Mixed Bank	0.015	-0.007	-0.234	-0.854	0.011	0.019
(Transactional Lending)	(0.023)	(0.015)	(0.234)	(0.836)	(0.013)	(0.019)
Corporate Loan × Mixed Bank	-0.013	-0.003	0.310	0.927	-0.007	-0.004
(Relationship Lending)	(0.032)	(0.012)	(0.316)	(0.936)	(0.014)	(0.013)
Firm Characteristics						
Firm Location Yerevan	-0.002		-0.235		-0.013	
	(0.022)		(0.195)		(0.011)	
Wholesale Retail Trade Firm	0.018		0.382		0.006	
	(0.041)		(0.306)		(0.017)	
Other Fields of Service Firm	-0.018		-0.121		-0.005	
	(0.024)		(0.104)		(0.010)	
Private Firm	-0.012		0.197		-0.007	
	(0.022)		(0.176)		(0.009)	
Fixed Effects						
Firm Fixed Effects	No	Yes	No	Yes	No	Yes
Constant	0.061**	0.051***	0.273	0.537	0.035***	0.015**
	(0.025)	(0.007)	(0.213)	(0.424)	(0.013)	(0.006)
R-squared	0.018	0.446	0.036	0.164	0.018	0.291
Observations	1,193	1,193	1,193	1,193	1,193	1,193

Table 2.10: Robustness Tests of Relationship Lending and Loan Performance

The Table reports robustness regression results for a sub-sample (or selection thereof) of 10,656 loan-time observations of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. The dependent variable is Loan Performance$_{ijkt}$ that equals one when a loan is in arrears for more than 90 days unless otherwise noted. For each robustness test, I rerun specifications of Table 2.3 in Columns (1) and (2) with firm characteristics or with firm fixed effects. To consume space, I only report the coefficient of "SME Loan × Mixed Bank" which measures the performance of transaction-based relative to relationship-based loans, i.e., SME loans of *Mixed* versus *Relationship Banks* (the reference group), as well as its standard error and the number of observations. Panel A reports regression results for alternative explanations of the main results: "Alternative relationship lending" defines the importance of relationship lending relative to the importance of other lending technologies and reduces the sample to *Relationship Banks* for which the relative importance of relationship lending stays constant for SME and corporate loans (4 out of 9 *Relationship Banks*) and *Mixed Banks* for which the relative importance of relationship lending is higher for corporate loans compared to SME loans (all 5 *Mixed Banks*), and "Relationship variables" adds relationship variables such as relationship duration, the scope of a firm-bank relationship and a dummy variable that equals one if a firm's debt exposure to a bank is above 50% and zero otherwise; "Full/Opposite sample" estimates specification (2) with just loan and firm characteristics for the full sample of 6,649 firms and 19,332 loans in the column Specification (2) and for opposite sample of 6,028 firms that receive 14,891 loans from either *Relationship* or *Mixed Banks* only in the column Specification (3). Panel B presents alternative non-performance measures: "Non-Performance" that equals one if a loan has any overdue payments on the principal amount and interest rate and zero otherwise and "Non-Performance 180 days" that equals one if a loan is delinquent for more than 180 days and zero otherwise. Panel C reports results for different loan characteristics. "Local currency loans" and "US dollar loans" uses only loans denominated in local currency or US dollars; "Loans after January 2011" uses only loans after January 2011; "W/o loans 50% around threshold" excludes loans 50% around the threshold; "Timing w/o SME & corp. Loans" excludes SME and corporate loans once a firm switches from SME to corporate loans for a *Mixed Bank*, meaning a switch from transactional to relationship lending; "Timing w/o SME Loans" excludes only SME loans once a firm switches from SME to corporate loans for a *Mixed Bank*. Panel D investigates different firm characteristics: "Trade, manufacturing, construction" uses only observations of firms from the trade, manufacturing or construction industries, while "Other fields of services" and "Other industries" uses only observations of firms either from other fields of service industries or remaining other industries; "New Customers" and "Old Customers" splits the benchmark regression into first time customers and repeated customers of a bank. Panel D examines different firm characteristics. Panel E relies on different estimation methods of the benchmark regressions. Row "Logit model" uses a logit model instead of a linear probability model for the benchmark specifications and reports odds ratios; "Matching on firm and loan amount" estimates the average differences in delinquencies for less than 90 days between SME loans of *Relationship* and *Mixed Banks* using a sub-sample of loans of the same firm with similar loan amounts: "Matching on firm and loan char." additionally matches on collateral, interest rate and rating. Specifications (2) and (3) correspond to a 0.5 and 0.25 standard deviation radius for each of the matching variables; the last two rows "Bank clusters" and "Bank×time and firm clusters" cluster standard errors at bank level only and use double clusters at bank-time and firm level, respectively. Definitions of the variables can be found in Table A.1 of Appendix A. If not otherwise noted, standard errors are clustered at firm level and presented in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10%.

Table 2.10 (continued): Robustness Tests of Relationship Lending and Loan Performance

	Specification (1) with firm characteristics			Specification (2) with firm fixed effects		
Variable Names	Coeff.	Std. Error	Obs.	Coeff.	Std. Error	Obs.
Panel A: Alternative Explanations						
Alternative relationship lending	-0.023**	(0.010)	5,898	-0.017*	(0.010)	5,898
Relationship variables	-0.026***	(0.009)	10,656	-0.026***	(0.008)	10,656
Full/Opposite sample	-0.020***	(0.004)	53,780	-0.016***	(0.005)	43,124
Panel B: Alternative Measure of Loan Performance						
Non-Performance	-0.005	(0.007)	10,656	-0.013**	(0.005)	10,656
Non-Performance 180 days	-0.007	(0.005)	10,656	-0.009**	(0.004)	10,656
Panel C: Loan Characteristics						
Local currency loans	-0.021*	(0.012)	5,946	-0.010	(0.011)	5,946
US dollar loans	-0.033***	(0.010)	4,710	-0.040***	(0.014)	4,710
Loans between January 2009-2011	-0.027**	(0.011)	2,881	-0.020	(0.015)	2,881
Loans after January 2011	-0.025**	(0.011)	7,775	-0.022**	(0.011)	7,775
W/o loans 50% around threshold	-0.026***	(0.010)	9,161	-0.025***	(0.009)	9,161
Timing w/o SME & corp. loans	-0.025***	(0.009)	9,885	-0.025***	(0.008)	9,885
Timing w/o SME loans	-0.025***	(0.009)	10,044	-0.025***	(0.008)	10,044
Panel D: Firm Characteristics						
Trade, manufacturing, construction	-0.007	(0.017)	3,367	-0.014	(0.013)	3,367
Other fields of services	-0.036***	(0.012)	5,715	-0.033***	(0.012)	5,715
Other industries	-0.022	(0.017)	1,574	-0.020	(0.016)	1,574
New customers	-0.020	(0.014)	3,389	-0.025*	(0.014)	3,389
Old customers	-0.027**	(0.011)	7,267	-0.027***	(0.010)	7,267
Panel F: Alternative Estimation						
Logit model	0.568**	(0.128)	10,656	0.444***	(0.105)	2,459
Matching on firm and loan amount	-0.026**	(0.010)	19,207	-0.026**	(0.013)	13,615
Matching on firm and loan char.	-0.043**	(0.020)	2,064	-0.062	(0.041)	404
Bank clusters	-0.025	(0.018)	10,656	-0.025***	(0.008)	10,656
Bank×time and firm clusters	-0.025**	(0.012)	10,656	-0.025***	(0.009)	10,656

Table 2.11: Relationship Lending and Information Use

The Table reports regression results of the variance equation of a multiplicative heteroskedaticity model based on Harvey (1976) and Cerqueiro et al. (2011) for a sub-sample of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. The model estimates the determinants of the mean and the residual variance of the Loan Spread$_{ijkt}$, defined as the loan interest rate minus the refinancing rate of the Armenian banks with the Armenian Central Bank in the upper and lower panels. The main independent variable is "Mixed Bank" which measures the effect of transactional lending on the residual variance in interest rates relative to relationship lending, i.e., SME loans of *Mixed* versus *Relationship Banks* (the reference group). Columns (1) reports results with loan characteristics and firm characteristics in the mean equation. In Columns (2)-(5) firm fixed effects are added to the mean equation. Column (3) adds loan origination fixed effects, Column (4) bank characteristics, and Column (5) other loan characteristics to the mean equation. The last row uses a Walt test to test whether the difference in coefficients when both banks rely on relationship lending, i.e., corporate loans of *Relationship* versus *Mixed Banks*, equals zero. Definitions of the variables can be found in Appendix A, Table A.1. Standard errors are clustered at firm level and presented in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10%.

Variance Equation	(1)	(2)	(3)	(4)	(5)
Corporate Loan	-0.342**	-0.109	-0.016	-0.016	-0.175
(Relationship Lending)	(0.167)	(0.174)	(0.176)	(0.175)	(0.170)
Mixed Bank	-0.248*	-0.696***	-0.416**	-0.665***	-0.741***
(Transactional Lending)	(0.131)	(0.157)	(0.178)	(0.153)	(0.152)
Corporate Loan × Mixed Bank	0.137	-0.088	-0.222	-0.200	-0.061
(Relationship Lending)	(0.197)	(0.258)	(0.283)	(0.261)	(0.269)
Variables in Loan Equation					
Loan Characteristics	Yes	Yes	Yes	Yes	Yes
Firm Characteristics	Yes	No	No	No	No
Bank Characteristics	No	No	No	Yes	No
Other Loan Characteristics	No	No	No	No	Yes
Firm Fixed Effects	No	Yes	Yes	Yes	Yes
Loan Origination Fixed Effects	No	No	Yes	No	No
Constant	2.490***	2.040***	1.221***	1.995***	1.978***
	(0.095)	(0.103)	(0.118)	(0.101)	(0.093)
Pseudo R-squared	0.087	0.204	0.323	0.208	0.221
VWLS R-squared	0.396	0.699	0.841	0.705	0.727
Observations (Loan Level)	4,441	4,441	4,441	4441	4,441
Corporate Loan = Corporate Loan × Mixed Bank	0.126	0.950	0.618	0.606	0.744

Figures

Figure 2.1: Identification strategy

The Figure illustrates the identification strategy for the main analysis, presenting bank types, loan types and corresponding lending techniques. The two bank types are *Relationship Banks* (control group) and *Mixed Banks* (treatment group). Each bank reports the importance of relationship lending by SME and corporate loan separately. *Relationship Banks* rely on relationship lending for both loan types, while *Mixed Banks* use relationship lending only for corporate loans and transactional lending (based on fundamental/cash flow analysis and collateral) for SME loans. The arrows indicate that loan performance of the *same firm* is compared for SME loans of the two bank types, when lending techniques differ (relationship vs. transactional lending), and for corporate loans, when both rely on relationship lending.

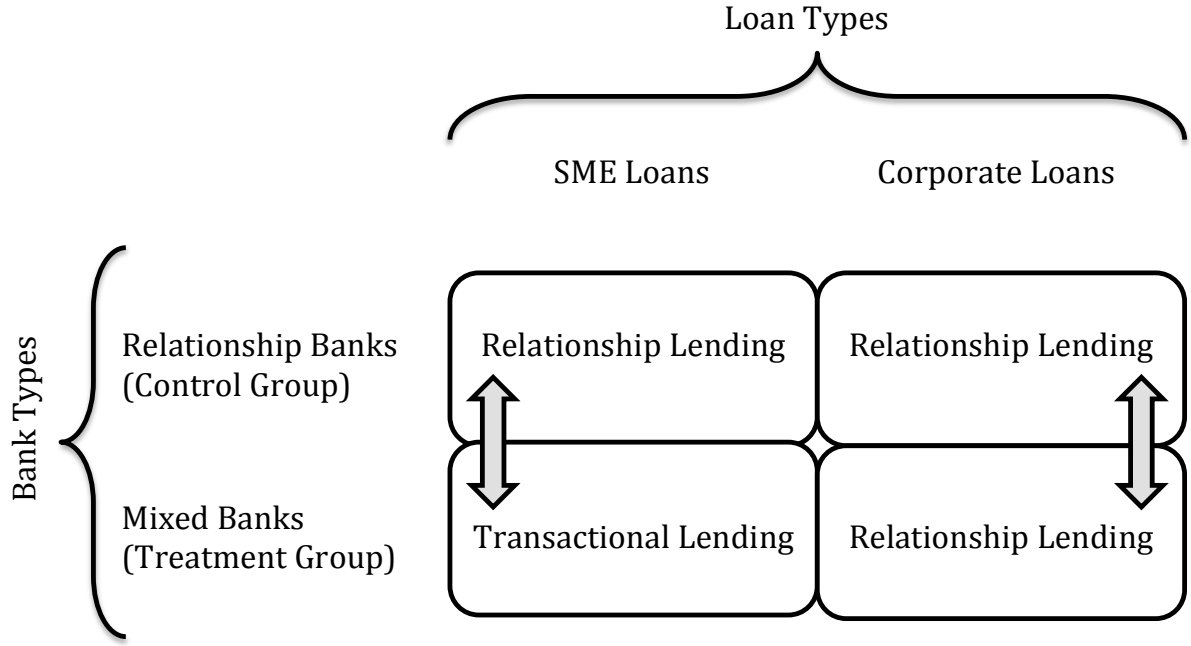
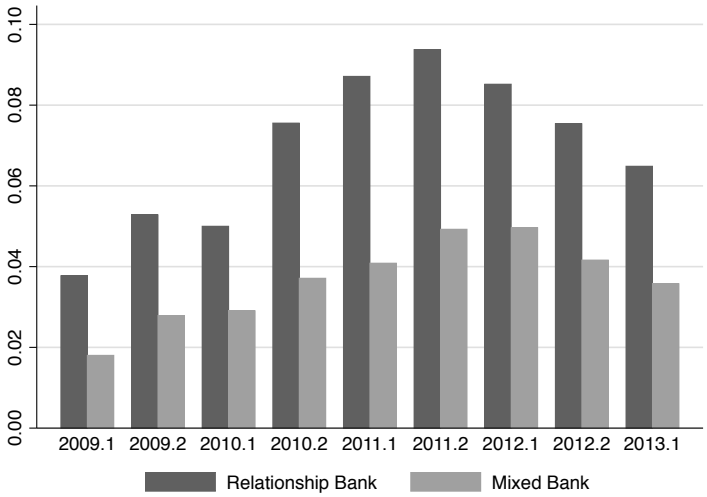


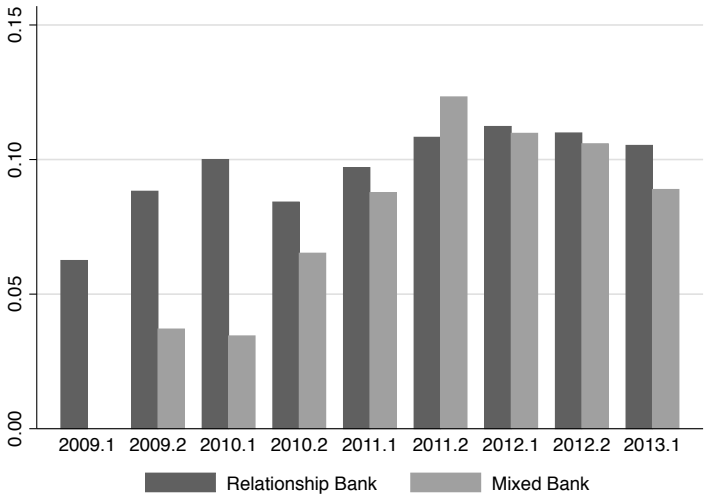
Figure 2.2: Loan Performance over Time for Sub-sample

The Figure shows the average percentage of loans in delinquency over 90 days for a sub-sample of 4,441 loans to 621 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. The upper figure shows average non-performance over time for SME loans for both *Relationship* and *Mixed Banks* (relationship vs. transactional lending), while the lower figure shows average non-performance for corporate loans for both bank types (both relationship lending).

(a) SME Loans (Relationship vs. Transactional Lending)



(b) Corporate Loans (Relationship vs. Relationship Lending)



A Variable Definitions

Table A.1: Variable Definitions

The Table reports variable definitions.

Variable Names	Definitions
Loan Performance	
Non-Performance	= 1 if a loan has any overdue payments on the principal and interest rate amount or overdue days, and = 0 otherwise.
Non-Performance 0-90 days	= 1 if a loan is non-performing for less than 90 days, and = 0 otherwise.
Non-Performance 90 days	= 1 if a loan is non-performing for more than 90 days, and = 0 otherwise.
Non-Performance 180 days	= 1 if a loan is non-performing for more than 180 days, and = 0 otherwise.
Default (loss/written-off)	= 1 if a loan has a loss/written-off status, and = 0 otherwise.
Loan Characteristics	
SME Loan	= 1 if a loan is classified as an SME loan based on the loan amount definition of a bank, and = 0 otherwise.
Corporate Loan	= 1 if a loan is classified as a corporate loan based on the loan amount definition of a bank, and = 0 otherwise.
Credit Classification	Credit classification of a loan (1 (worst rating) and 5 (best rating)).
Interest Rate	Annual contractual interest rate at loan origination.
Loan Spread	Loan interest rate minus the refinancing rate of the Central Bank of Armenia.
Loan Amount in US$	Loan amount at loan origination in US dollars.
Collateral	= 1 if collateral was pledged at loan origination, and = 0 otherwise.
Guarantee	= 1 if a guarantee was given at loan origination, and = 0 otherwise.
Loan Maturity in Months	Number of months between loan origination and maturity.
Other Loan Characteristics	
Loan Location in Yerevan	= 1 if the location of the loan is in Yerevan, and = 0 otherwise.
Wholesale and Retail Trade Loan	= 1 if the industry of the loan is in the whole and retail trade industry, and = 0 otherwise.
Other Fields of Service Loan	= 1 if the industry of the loan is in other fields of the service industry, and = 0 otherwise.
Loan in USD	= 1 if the currency denomination of the loan is in USD, = 0 otherwise.
Relationship Characteristics	
Relationship in Months	Duration of a bank-firm relationship in months.
Scope	= 1 if the firm has additional products (e.g., credit lines, leasing, factoring, overdrafts) with a bank, and = 0 otherwise.
Primary Bank	=1 if more than 50% of a firm's outstanding debt is originated by one bank, and = 0 otherwise.
Number of Relationships	Number of banks with which a firm has outstanding loans.
Multiple Relationships	= 1 if the firm has outstanding loans from multiple banks, and = 0 otherwise.
Firm Characteristics	
Firm Location Yerevan	= 1 if the location of the firm is in Yerevan, and = 0 otherwise.
Wholesale and Retail Trade Firm	= 1 if the industry of the firm is in the whole and retail trade industry, and = 0 otherwise.
Other Fields of Service Firm	= 1 if the industry of the firm is in other fields of the service industry, and = 0 otherwise.
Private Firm	= 1 if the firm is a private firm, and = 0 otherwise.
Bank Characteristics	
Relationship Bank	= 1 if a bank reports a high importance (frequency of use) of relationship lending for SME and corporate loans, and = 0 otherwise.
Mixed Bank	= 1 if a bank reports a high importance (frequency of use) of relationship lending for corporate loans but less importance for SME loans, and = 0 otherwise.
Large Bank	= 1 if the Central Bank of Armenia considered the bank to be large in terms of total assets at the of 2009. Alternatively, banks are classified as medium-sized and small.
Foreign Bank	= 1 if more that 50% of its equity is foreign-owned (excluding investors with Armenian origin).
High Average Hierarchy	= 1 if the number of layers a loan has to pass to for approval is above 2.

B Test for Discontinuity at the Threshold

In this section, I examine the distribution of loans around the threshold that determines a loan to be an SME or corporate loan. A natural question that arises is whether banks or firms are manipulating loan amounts in order to give out or receive either SME or corporate loans. *Mixed Banks*, for example, could intentionally give out loans with loan amounts just below the threshold in order to avoid giving out a corporate loan that might be associated with higher costs since relationship lending becomes more important. Likewise, firms could apply for loans just below the threshold in order to circumvent possibly higher screening and monitoring activities of banks.

In general, only complete manipulation but not partial manipulation results in identification problems. While complete manipulation assumes that the assignment rule is under complete control of agents, partial manipulation occurs when agents can only partially influence the assignment rule and the rest remains idiosyncratic (McCrary (2008), p 700).[66] In the present case, threshold definitions are not publicly known and differ across banks in amount and currency. Half of the banks set thresholds in USD, while the rest sets it in AMD. At the same time, loans are issued in different currencies. For firms, it is more difficult to manipulate their loan amounts, as they are less likely to know the exact thresholds for each bank. At banks, loan officers might be able to manipulate loan amounts. Exchange rate fluctuations might, however, still add an idiosyncratic component (see Garmaise and Natividad (2014) for similar ideas), suggesting partial manipulation.

Even if complete manipulation occurs, it should not influence the main results, since the identification comes from loans further away from the threshold and not around the threshold. Most likely, for *Mixed Banks* lending techniques do not just switch from transactional to relationship lending once a loan passes the loan amount threshold but rather get less transaction-based and more relationship-based with the loan size. Therefore, the further away a loan is from the threshold the more prominent the difference in lending techniques will be for *Mixed Banks* across loan types and for SME loans between *Relationship* and *Mixed Banks*. In unreported results, I confirm that leaving out loans exactly at or around the threshold does not alter the main results.

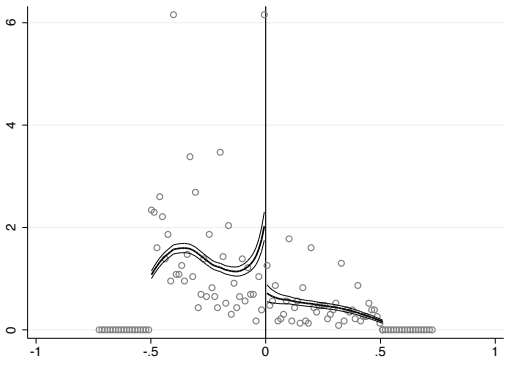
In order to formally check for manipulation around the threshold, I rely on a methodology developed by McCrary (2008) that tests for the discontinuity at the threshold in the density function of the running variable (loan amount threshold). The upper panel of Figure A.1 plots the density functions of loan amounts with the threshold normalized to zero and a range of 50% around the threshold for the total sample and the sub-sample. Both figures reveal a discontinuous jump at the threshold which is confirmed by coefficients of -1.06 (-1.46) and standard errors of 0.14 (0.28). Errors in the assignment of loans to SME and corporate loans might occur since banks do no explicitly specify whether the threshold is an upper or lower bound or might give approximate amounts. The lower panel of Figure A.1 plots the same density functions as above, leaving out loans exactly at the threshold. The discontinuous jump disappears with coefficients of 0.02 (-0.11) and standard errors of 0.14 (0.30).

[66] Van der Klaauw (2002), DiNardo and Lee (2004), and Lee (2008) present some examples of plausible partial manipulations that do not influence results.

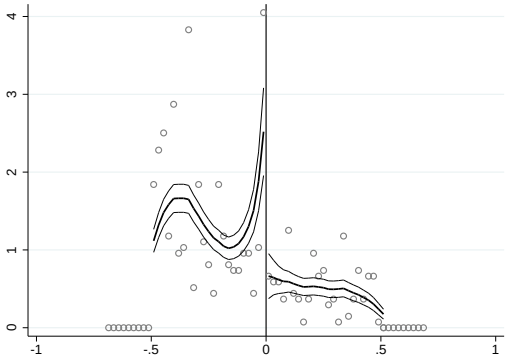
Figure A.1: Density of Loan Amounts around Threshold

The Figure shows the density of loan amounts at the threshold that is normalized to zero and in a range of 50% based on McCrary (2008). The upper figures show the density for the total sample of 19,332 loans to 6,649 firms and the sub-sample of 4,441 loans to 271 firms that received loans from both *Relationship* and *Mixed Banks* between January 2009 and June 2013. The lower figures repeat the analysis for the total and sub-sample, excluding loan amounts directly at the threshold.

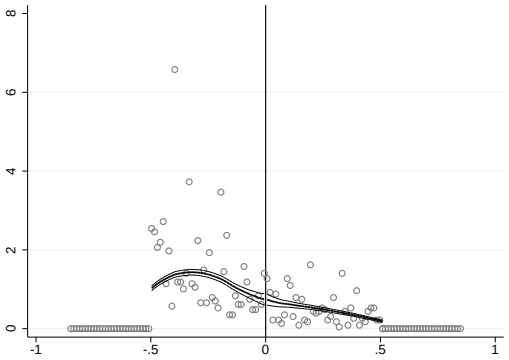
(a) Total Sample



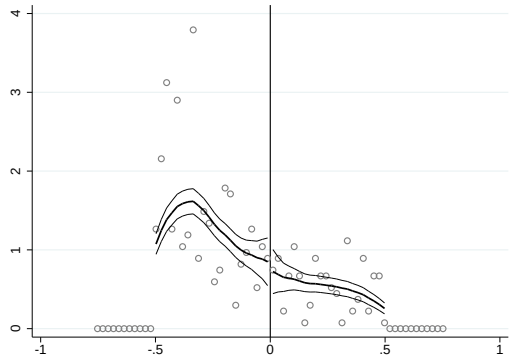
(b) Sub-sample



(c) Total Sample



(d) Sub-sample



C Regression Model with Multiplicative Heteroskedasticity

The regression model with multiplicative heteroscedasticity based on Harvey (1976) is defined as:

$$y_i = \beta' X_i + u_i, \tag{A.1}$$

$$Log(\sigma_i^2) = \gamma' Z, \tag{A.2}$$

where (1) is the mean equation and (2) the variance equation. The identifying assumptions for the model are:

$$E(u_i|X_i) = 0, \tag{A.3}$$

$$E^2(u_i|Z_i) \equiv \sigma_i^2 = \exp(\gamma' Z_i), \tag{A.4}$$

where y_i is the depending variable, X_ia vector of explanatory variables in the mean equation, u_i is a disturbance term, σ_i^2 the residual variance, and Z_i a vector of explanatory variables in the variance equation.

Under the normality assumption, the conditional distribution of y_i is given by:

$$y_i|X_i, Z_i \overset{d}{\rightarrow} N(\beta' X_i, \exp(\gamma' Z_i)), \tag{A.5}$$

The heteroscedastic regression model is estimated with Maximum-Likelihood (MLE) by maximizing the following log-likelihood with respect to β and γ:

$$LogL = \frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\gamma' Z_i - \frac{1}{2}\sum_{i=1}^{n}\exp(-\gamma' Z_i)(y_i - \beta' X_i)^2 \tag{A.6}$$

Harvey (1976) shows that this approach is analogous to estimating the mean Eq. (1), and taking the squared-residuals as the raw estimates of the individual variances, which are subsequently used to estimate Eq. (2). This two-step approach leads to a substantial loss of efficiency vis-à-vis the MLE.

Chapter 3

An Anatomy of Central and Eastern European Equity Markets

Coauthors: Lieven Baele, Geert Bekaert

3.1 Introduction

Following a liberalization process in the late 80s and early 90s, emerging equity markets have become an integral part of global equity portfolios. Their properties have, not surprisingly, been studied in detail (see, e.g., Harvey (1995), Rouwenhorst (1999b) and more recently, Bekaert and Harvey (2014)). Much less is known about emerging markets in Eastern and Central Europe, which liberalized typically later, following the fall of the Iron Curtain in 1990 and subsequent liberalization and privatization waves in a number of these countries. Previous papers typically analyze only the largest markets (Russia, Hungary, Poland, Czech Republic), mainly because commercial indices (e.g., MSCI, Datastream, and S&P 500) are only available over a sufficient time span for these particular markets (see Chelley-Steeley (2005), Schotman and Zalewska (2006), Gilmore et al. (2008), Savva and Aslanidis (2010b), Caporale and Spagnolo (2011), Syllignakis and Kouretas (2011), and Gjika and Horvath (2012) amongst others). In this article, we provide a comprehensive and detailed overview of Central and Eastern European (CEE) equity markets from the mid-1990s on until now and evaluate the value of investing in these markets for global investors.

With the aim of having comparable indices for longer time spans and a larger cross-section of equity markets, we build our own indices based on a sample of more than 2,000 individual stocks from 14 CEE equity markets. These stocks pass a series of inclusion tests on data availability and liquidity. After careful checks of the return data, we construct tailor-made value-weighted indices that target about 85% of the total market capitalization as well as weekly returns in excess of the 3-month T-bill rate, both at the country and aggregate CEE level.

For nearly all countries, our self-constructed indices start (often substantially) before the earliest starting date of the commercial indices. Looking at overlapping periods, we observe considerable differences between our tailor-made indices and all of the commercial ones. We find correlations of our indices with the commercial ones to be substantially below 1, especially for the smaller equity markets. Quite strikingly, we find equally low correlations among the different commercial indices. We show that differences between our and commercial indices are not driven by using market rather than free-float adjusted weights.

Next, we characterize and evaluate equity market development in the various CEE markets. Considerable cross-sectional differences exist in the level of market development, as measured by market capitalization over GDP, liquidity proxies, and concentration indices. Russia, Turkey, and Hungary are the most developed markets, while Kazakhstan, Latvia, and the Slovak Republic are the least developed ones. We also examine which economic, financial and institutional factors determine market development, finding that market specific laws and reforms, in particular, the implementation of insider trading laws, as well as liberalization, affect market development overall, while institutional and political reforms as well as economic and financial openness foster equity market development within a country. We will explore the drivers of market integration in much more detail in a subsequent paper.

The CEE markets are sufficiently developed to provide another avenue for the diversification of global equity portfolios. We therefore examine how their correlations relative to the world market and other emerging markets have evolved over time. We generally find, as is true for emerging markets more generally, an upward trend in correlations which is due to both increasing global market exposures and a reduction in country-specific risks. Especially the larger, more developed equity markets, have correlations with developed markets of nearly 80%, and

hence offer little diversification potential. Many of the smaller markets, typically categorized as frontier markets, have correlations below 50%, and do offer scope for diversification. Increasing benchmark betas are a stronger source of increasing correlations for those markets than are decreasing country-specific volatilities. Based on the model by Heston and Rouwenhorst (1994), we show that despite increasing industry effects, country factors still dominate in CEE markets, suggesting that country diversification is still valuable. When exploring the cross-section of correlation dynamics, we find cross-listed, large, and liquid stocks to have higher benchmark correlations and betas. We do observe, however, a catching-up effect of especially the "home-only" and small stocks over the final 10 years of our sample.

We also verify whether the CEE equity markets exhibit various cross-sectional properties of individual returns, documented for developed markets (Hou et al. (2014)) and emerging markets (see Rouwenhorst (1999a)). We confirm the presence of a size, value, low volatility, and (il)liquidity effect in CEE markets, but do not find evidence for a momentum or a low beta effect. These results continue to hold when we take free-float adjustments into consideration. We show how these premiums can be jointly exploited using the parametric portfolio policy methodology of Brandt et al. (2009) that decomposes portfolio weights into a benchmark weight and an active component that over- or underweights stocks based on their characteristics. The optimal strategy that accounts for short-sale constraints has a bias towards small firms, value firms and past winners (positive momentum). We find the size effect to dominate the value effect, which in turn dominates the momentum effect. When we add volatility, illiquidity, and local beta as additional characteristics, the optimized portfolio is tilted towards low volatility and less liquid stocks and, surprisingly, towards high, not low, beta stocks. The optimized portfolio with three characteristics has an annualized alpha with respect to CEE benchmark index of 13.9% (but also a relatively low beta and volatility). The alpha increases with an additional 2% when all 6 characteristics are taken into account.

The rest of the chapter is structured as follows. Section 3.2 offers a detailed description of equity indices we construct and analyzes main properties relative to existing commercial indices. Section 3.3 explores market development and liquidity indicators over time. Section 3.4 analyzes time-varying co-movements and correlations, while Section 3.5 focuses on the cross-section of

expected returns. Section 3.6 concludes.

3.2 Data Description

3.2.1 Data Sources and Stock Selection

Our indices and market statistics are based on firm level data for 16 equity markets, namely Bulgaria, Croatia, the Czech Republic, Estonia, Hungary, Kazakhstan, Latvia, Lithuania, Poland, Romania, Russia, Serbia, the Slovak Republic, Slovenia, Turkey, and Ukraine. From Datastream, we download data for all stocks (including American Depositary Receipts and Global Depositary Receipts) associated with any of the 16 country codes, yielding a total of 8,686 stocks over the period January 1990 up to May 2011. To avoid a look-ahead bias, we also include dead or suspended stocks.[67] For each stock, we download at the daily frequency the price and return index, both in local currency and in US dollars, volume (total number of traded shares), market capitalization, both in local currency and in US dollars, and the industry code.

We apply the following filters to the raw data. First, we eliminate all stocks that have no data. This rather drastically reduces our sample from 8,686 to 5,356 stocks. Second, we require a stock to be sufficiently liquid. As for many firms the data on trading volume is often missing or not reliable, we use the percentage of daily zero price returns in local currency within a year to eliminate illiquid stocks (see, e.g., Lesmond et al. (1999) and Lesmond (2005)). We consider a stock to be sufficiently liquid if in 25 percent of all trading days within the last year the daily price return in local currency was different from zero.[68] Bekaert et al. (2007b) show that this measure is highly correlated with more common (high-frequency) liquidity measures. We include a firm in our sample if it meets our liquidity criterion for at least one year over the full sample. The liquidity criterion reduces the sample further from 5,356 to 2,510 stocks. Third, we removed all preferred shares from our sample as they clearly showed bond rather

[67] For delisted and dead stocks, we put all data after the exact delisting or bankruptcy date to missing. If such a date is not given, we take the last observation after which no change in the return index in local currency appeared any more as the delisting or bankruptcy date and again set all values after that date to missing. If the return index did not change for more than a year before the delisting or bankruptcy date, we take the last date on which a change in the return index appeared as the delisting or bankruptcy date.

[68] For Bulgaria and the Slovak Republic we use 24 percent of all trading days as it significantly changes the number of liquid stocks and the starting date.

than equity-like properties.[69] In a last step, we compare the liquidity of cross-listed stocks, ADRs, and GDRs with the respective home stock for each year by calculating the percentage of non-zero price returns in local currency. When the foreign listing is more liquid, we use its price and total returns instead of the local data.[70] In many CEE markets, with Russia as an important example, trading is much more active in the foreign listed stocks. Our volume data count both the volume in the home and foreign market(s) (in US dollars). Our final sample consists of 2,131 stocks. Notice that we do not have a specific size threshold, as we see no ex-ante reason to exclude small stocks. Of course, many of the illiquid stocks also tend to be small.

Table 3.1 shows the composition of local common, local preferred and foreign stocks (cross-listed, ADR, GDR) for the total sample and for the liquid sample, distinguishing between active and dead stocks. As many stocks in CEE markets are highly illiquid, the total number of stocks dramatically reduces to often less than half of the total sample of stocks for the liquid sample. The number of selected stocks differs greatly between countries, ranging from 8 for Kazakhstan and 14 for the Slovak Republic to 397 for Turkey and 475 for Poland. While in most countries the number of foreign stocks is small, we observe many cross-listed stocks in Poland (20), Estonia (25), the Czech Republic (37), Turkey (57), Hungary (66), and Russia (167) corresponding to respectively 4%, 47%, 17%, 13%, 47%, 33% of the total number of liquid stocks. We analyze these cross-listed stocks in more detail in Section 3.4.4.

The last two columns of Table 3.1 show the average firm and market size in millions of US dollars in 2010. The average firm market capitalization is calculated as the average market capitalization across selected firms per day and across time. The average total market capitalization stands for the total market capitalization of the selected firms per day and across time. While the Czech Republic, Russia, and Kazakhstan have the largest firms on average ($6.2bil, $4.3bil, and $2.7bil), not surprisingly Russia, Turkey and Poland represent the largest markets

[69]To decide whether a preferred stock is bond or equity like, we calculate the ratio of the variance of the common stock over the variance of the preferred stock. If that ratio is lower than 1.2, signaling that the volatility of the preferred stock is of about the same magnitude as that of the common stock, we conclude that it is equity like, and include it in our sample. If instead the ratio is higher than 1.2, signaling that the volatility of the common stock is much higher than that of the preferred stock, we categorize the preferred stock as bond-like, and exclude it from the sample.

[70]Some foreign stocks do not have a home stock but are still included in our sample if they meet the liquidity criterion.

in the region with market capitalizations well over $100 billion.

3.2.2 Index Construction

Based on this sample of firms, we construct our own daily country indices. We update the composition of the index on a yearly basis (first trading day of the calendar year). For a firm to be included in the index, it needs to be active and have passed the liquidity criterion on January 1. Using returns including dividends and market capitalization in U.S. dollars, we construct value-weighted and equally-weighted indices with an objective of meeting 99%, 85% or 70% of total market capitalization.[71] For a given country, we only start constructing indices from the moment that there are at least three liquid stocks (MSCI follows a similar procedure for emerging markets, see MSCI (2011)). Finally, before constructing our indices, we check for extreme daily returns, defined as an absolute return larger than 200 percent. As a first step, we try to confirm the extreme return from alternative sources, such as Bloomberg, Yahoo Finance, the local stock exchange or the website of the firm (see, e.g., de Groot et al. (2012)). If we cannot confirm the extreme return, we replace it by the value that is most frequently reported across the different providers. In most cases, this happens to be the lowest absolute return across the different sources.[72]

3.2.3 Preliminary Return Analysis

Table 3.2 reports summary statistics for the weekly excess returns (including dividends) of our tailor-made value-weighted index, both at the country and aggregate CEE level, as well as benchmark returns from MSCI CEE, Emerging Markets, and Frontier Markets for an overlapping sample period between January 2002 and May 2011. For the Ukraine and MSCI frontier markets, the actual starting date is later than 2002, and Kazakhstan and Serbia are excluded due to their short time series. As is common for commercial indices, we focus on an index

[71] In practice, it is not possible to achieve the size-based segments exactly. Therefore, we introduce market capitalization size ranges of $70\% \pm 5\%$, $85\% \pm 5\%$, and $99\% + 1\%$ or -0.5%, similar to S&P's methodology (MSCI (2011)).

[72] We removed another 6 stocks (4 from Russia, and 1 each from Slovenia and Ukraine) because they had too many outliers. Because the weight of each of the removed stocks in the total market capitalization is (much) smaller than 0.1 percent, index values are not affected.

that targets about 85 percent of total market capitalization.[73] Weekly returns are based on daily index values observed on Wednesdays. All returns are in US dollars and in excess of the 3-month T-bill rate.

The first column shows the starting month and year of our indices. We observe the longest time series for Turkey (as from 01-91), Hungary (as from 01-92) and Poland (01-93). Except for Ukraine (as of 01-06), all countries have at least 11 years of daily returns. Annualized returns vary widely across markets but are mostly above 10% except for Latvia, Ukraine, and MSCI Frontier Markets which can be explained by the shorter sample periods of the latter two. The annualized value-weighted returns range between 7% for Ukraine to 44% for Bulgaria, and their standard deviations from 24% for Slovenia to 41% for Turkey. Returns exhibit on average slightly negative skewness, substantial excess kurtosis, but relatively low autocorrelation. Compared to the MSCI Emerging Market and Frontier Market indices, most CEE countries have higher means and standard deviations, while other statistics are relatively similar.

In the last column, we conduct a variance ratio test of the null that the variance of monthly returns (4 weeks) over the variance of weekly returns is equal to 1 or put differently a test of standard random walk null. Following Campbell et al. (1997), the q-period variance ratio test statistics is defined as $VR(q) \equiv \frac{Var[r_t(q)]}{qVar[r_t]}$, where $r_t(q) = r_t + r_{t-1} + ... + r_{t-q}$. The distribution of the test statistic uses a heteroskedasticity consistent asymptotic variance matrix. In our application, we set q equal to 4 (weeks). We reject the null in only 6 cases, with the rejections being at the 1% significance level for Bulgaria, Estonia, and the MSCI Frontier market index. The ratio is above one for all but one country, indicating positive return autocorrelations.

3.2.4 Comparison with Commercial Indices

Table 3.3 provides a comparison of our tailor-made indices to alternative benchmark indices for overlapping time periods. We consider MSCI, Datastream and S&P as benchmark providers. For each country, we report the starting dates of the benchmark index, the slope coefficient β from a regression of our index returns on the benchmark index returns and pairwise correlations.

[73] Results for other thresholds (70%, 99%) as well as for equally rather than value-weight indices (70%, 85%, 90%) are available upon request.

In the row below, we report p-values from tests of the hypothesis that betas and correlations, respectively, are equal to one. The correlation (ρ) test uses the Fisher transformation [$z = (1/2)\ln((1+\rho)/(1-\rho))$] for which the asymptotic standard error is $1/\sqrt{T-3}$.[74] The final columns report correlations between the returns on the alternative benchmark indices, and the corresponding p-values of a test that they are equal to one. All statistics are based on total weekly dollar returns in excess of the 3-month T-Bill rate.

A first striking observation from Table 3.3 is that there is wide dispersion in starting dates. There is no provider that systematically has the longest time series, even though for most countries S&P is among the first to report an index. While S&P and MSCI have indices for nearly all countries (MSCI misses Latvia and Slovakia), country representation is more limited for Datastream (no indices for Croatia, Slovakia, Ukraine, and the 3 Baltic states). Our self-constructed indices have in the majority of cases a starting date that is at least as early as the earliest commercial index. The exception are the S&P indices starting (several) years earlier for Bulgaria (6 years), Lithuania (3 years), Slovakia (4 years), and Ukraine (8 years). The later starting dates of our indices for these countries reflects our requirement of at least 3 sufficiently liquid stocks at a daily frequency before starting the index.

Second, the betas and correlations of our index with the alternative benchmark indices are not always close to 1 and not necessarily close to one another. For larger countries such as the Czech Republic, Hungary, Poland, Romania, Russia and Turkey the betas and correlations are mostly above 0.9. For smaller countries such as the Slovak Republic and the Ukraine, however, the betas and correlations are between 0.14 and 0.20. We overwhelmingly reject that betas are equal to 1. For the larger countries such as Czech Republic, Hungary, Poland, Romania, Russia and Slovenia, we cannot reject that betas are equal to 1 for the correlation of our tailor-made indices and the Datastream indices (sometimes also for the MSCI indices). For correlations, the test also overwhelmingly rejects that correlations are equal to 1, except for the correlation of Romania's tailor-made index with the corresponding Datastream index. On average, betas (correlations) range from 0.61 (0.63) for the S&P to 0.92 (0.90) for Datastream across countries. Note that the country coverage plays a role here as well, as it is widest for

[74] Since for the null $\rho = 1$ the Fisher transformation is not valid, we use $\rho = 1 \approx 0.99$ as an approximation.

the S&P and weakest for Datastream. At the regional level, when we compare our tailor-made CEE index to alternative benchmark indices, betas and correlations range from 0.91 to 1.10.

Third, moving to the next set of columns, the correlations among the alternative benchmark indices are reasonably close for the large countries, but can be quite low for other markets, e.g. less than 75% for Bulgaria, Estonia, Romania, Slovenia, and the Ukraine. Thus, similar differences prevail among benchmark providers. Interestingly, we overwhelmingly reject that correlations among different providers are equal to 1, except for the correlation of the Datastream and S&P index for Turkey.

Several elements of index construction can cause differences between the different commercial indices and our own, including selection criteria regarding size and liquidity, the handling of cross-listings and free float, and the reviewing period to include or exclude stocks. In terms of size criteria, we do not impose minimum market capitalization requirements, whereas all the other indices do. For example, stocks must have a float-adjusted market capitalization above US$ 1 billion as of the rebalancing reference date for Standard & Poor's (Standard & Poor's (2012)), meaning that only larger companies are included in the index. In terms of liquidity, we require that stocks must have had non-zero price returns for at least 25% of all trading days. MSCI relies on the percentage of annual value traded ratios (trading value over free float-adjusted market capitalization) over the past 3 and 12 months, while S&P requires stocks to have a 3-month average daily value traded above US$ 2 million as of the rebalancing date of the index.

When it comes to cross-listings, we replace home stocks with the most liquid ADRs, GDRs or direct cross-listings which is similar to the MSCI and S&P methodologies. Datastream ignores factors such as liquidity and cross-listings overall. With respect to exposure, we try to represent 85% of the total market capitalization which is similar to MSCI. S&P aims at a 80% market capitalization representation, while Datastream targets only 75% to 80% of the market. Reviewing periods vary from quarterly for the MSCI and Datastream indices to annual for our and the S&P indices. Most importantly, there are considerable differences in free-float adjustments. While Datastream and our index do not adjust for free-float, attempting to represent the full economic value of a market, MSCI and S&P instead adjust their indices

for free float. In particular, MSCI only includes stocks that have a free-float adjusted market capitalization equal or above 50% of the minimum size requirement, while S&P simply uses free-float adjusted market capitalizations and excludes stocks closely held by "strategic control" shareholders, other publicly traded companies and government agencies (Datastream (2008); MSCI (2011); Standard & Poor's (2012)).

Because government agencies still play a relatively large role in Central and Eastern European markets, the free float adjustment is perhaps the most important source of index differences and we investigate it in more detail. To this end, we build free-float adjusted indices and compare them to our proposed index as well as to alternative benchmark indices. Unfortunately, the information on free-float often becomes available much later than the original trading dates of stocks. In addition, on average, only 70% of stocks have data on free-float, ranging from 40% for Bulgaria up to 97% for Estonia (see Table 3.2). Nevertheless, we create two float adjusted indices from our data. The first uses only stocks for which the free-float adjusted information is available, but weights them using their full market capitalizations and we indicate it by FF1. The second index, denoted FF2, uses these stocks weighted by their actual float-adjusted market capitalizations. In Table 3.4, we characterize the differences between the original tailor made index (Original), its float-adjusted versions (FF1 and FF2) and the benchmark indices, by computing annualized tracking errors between various combinations of indices for overlapping time periods. The first set of columns compares the different versions of our tailor-made indices. The second set of columns report tracking errors of our original index relative to the benchmark indices from MSCI, Datastream, and S&P. The third set of columns replaces our original index by FF2, and the fourth set shows relative tracking errors among the currently available benchmark indices.

In the first set of columns, the first column suggests that focusing on the typically larger stocks for which a float adjustment can always be computed gives rise to sometimes substantial tracking errors. However, for the overall index, the tracking error is only 2.14%. Comparing the first with the second column reveals that implementing the actual float adjustments on the weights always increases the tracking error, on average by about 4%. The index tracking error is now 6%. The FF1/FF2 column measures a pure float adjustment effect as stock coverage is

identical. For most countries, float adjustment (column 3) produces larger tracking errors than the use of different stocks in the sample. The index has a tracking error of 5.6%.

The second set of columns shows that the tracking error of our CEE index is 7.14% with respect to the MSCI index, 9.56% with respect to the Datastream index and a quite large 17.19% relative to the S&P. The tracking errors for individual countries are often much larger still. To help interpret these numbers, the fourth set of columns looks at the relative tracking errors among the available benchmark indices, showing them to be of similar magnitude.

Nevertheless, the float adjustment is clearly not the main reason for the differences we observe between our indices and the available benchmark indices. Comparing the results in the second set of columns (original index versus benchmarks) with those in the third set of columns (float-adjusted indices versus benchmarks), the tracking error is larger for the float-adjusted indices than it is for our original index in at least 50% of the countries. It is substantially larger for the regional CEE index.

We conclude that different indices show sometimes large tracking errors, but that the tracking error of our index relative to the current benchmarks is not driven by float adjustment. We therefore opt to use non-float adjusted indices, which have longer and broader market coverage.

3.3 Market Development and Liquidity Indicators

This Section characterizes the market development in each of the 16 CEE equity markets by tracking the evolution of 5 key indicators. Our first measure, the ratio of total market capitalization over GDP, tracks the overall size of the equity market relative to the real economy. Measures (2) and (3), namely equity market turnover and the average percentage of non-zero daily returns, both track the evolution of market liquidity. Amihud and Mendelson (1986) show that turnover, calculated as the ratio of total dollar trading volume per year over the end-of-year market capitalization, is negatively related to illiquidity costs. Our preferred liquidity measure, however, is the average percentage of non-zero daily returns, calculated as the yearly average of the value-weighted share of non-zero daily price returns (in local currency). The main advantage of this measure is that it does not require detailed transaction data (such as bid-ask spreads), typically not widely available in emerging markets (see Bekaert et al. (2007b)

for a detailed discussion). Our last measures track concentration at the firm and industry level, measured using a Herfindahl (HHI) index and the share of the largest 3 firms (industries) (C3). We expect concentration to decrease as markets become more developed. Section 3.3.1 discusses the state of market development at the end of our sample in 2010. In Section 3.3.2, we perform panel regressions to better understand the cross-country determinants of market development over time.

3.3.1 Market Development in 2010

Table 3.5 reports the 2010 values for the different market development indicators for our 16 CEE markets. Where available, we provide comparable statistics for two large developed markets, namely Germany and the US. We also indicate whether each market is frontier or emerging according to the MSCI classification.

The first column of Table 3.5 shows that the majority of countries (and in particular frontier markets) have a ratio of market capitalization over GDP below 15 percent, which is far below levels observed in the US (118%) or Germany (43.3%). Exceptions include Russia (58.8%), Turkey (38.6%), Poland (30%) and Hungary and the Czech Republic (both around 20%). Markets with the highest level of market capitalization over GDP also tend to have the largest turnover ratios (column 2). Except for Turkey (136.8%), Hungary (107.3%) and Russia (81.3%), all markets have turnover ratios below 50%, which is considerably lower than in Germany (114%) and the US (176%). Half of our countries even have turnover ratios below 10%. The percentage of non-zero returns (column 3) is more than 90% in the Czech Republic, Hungary, Russia, and Croatia and never drops below 60%, in part because our firms satisfy a liquidity criterion, calculated using this measure. The concentration indicators imply that most markets are highly concentrated, both at the industry and the firm level and more concentrated than are Germany and the US. Countries that have relatively low industry concentration indices using the Herfindahl index include Slovenia, Bulgaria and Hungary, with Lithuania, and Poland joining Slovenia among the three least concentrated markets using the C3 index. Because the ranking based on the HHI versus C3 indices is quite similar, we use the HHI indices going forward.

To rank the different countries according to their level of market development, we follow an ordinal approach. First, we rank indicators that increase with development (market cap/GDP, turnover, percentage zero returns) from low to high, and those that decrease with development (the Herfindahl indices both at the industry and firm level) from high to low. Second, we replace each country-indicator observation by the country's rank number for that specific indicator. Third, we calculate a joint market-development indicator by taking the average across the different indicator ranks. The country with the highest indicator value gets rank 1, the one with the lowest 16. The second to last column of Table 3.5 reports the resulting rank. The top-3 countries are Russia, Turkey, and Hungary, respectively. Slovenia (5) is the highest ranked frontier market, followed by Lithuania (6) and Ukraine (8). The bottom three countries are Kazakhstan, Latvia, and the Slovak Republic.

Overall, it seems that there is great variety in development across Central and Eastern European markets with large markets such as Russia and Turkey catching up with developed markets and small markets such as Kazakhstan, Latvia, and the Slovak Republic being just at the beginning of the development process. To better understand equity market development over time we performed trend and structural break tests on each of the 5 market development indicators.[75] While both trend and break tests reveal improving market development indicators in most countries, the effects are mostly statistically insignificant and small in economic magnitude, and are therefore relegated to the online Appendix.

3.3.2 Drivers of Market Development

During the sample period, many countries undertook various reforms such as stock market liberalization or introduction of insider trading laws with implementation dates somewhat clustered but far from perfectly correlated across countries. Moreover, the reform process is not complete yet along some dimensions and for some countries, making it informative to know which reforms are associated with stock market development. We therefore relate our market development indicators to: (i) institutional and political reforms measured by the EBRD Transition

[75] In particular, we employ the linear time trend test by Bunzel and Vogelsang (2005) based on a simple time series model: $y_t = \beta_1 + \beta_2 t + u_t$, where y_t stands for the variable of interest and t for the linear time trend. We test for the null hypothesis of $\beta_2 = 0$ and use a Daniell kernel to estimate the error terms in order to maximize the power of the test in small samples. We report detailed estimation results in Table 2 of the online Appendix.

Indicator that assesses progress in transition based on a set of transition indicators, the Political Constraint Index, a political risk measure based on Henisz (2002), which increases when political actors face more vetos, and thus more political constraints, enacting policy changes which leads to less political risk, and the EU accession announcement year (EU Accession), (ii) capital controls and liberalization measures that include the intensity of capital controls indicator of Chinn and Ito (2006) and Chinn and Ito (2008) (Financial Openness), dummies for the official liberalization year (Official Liberalization) and for the first sign of liberalization year (First Sign of Liberalization) from Bekaert et al. (2005b), (iii) market specific laws and reforms that include dummies for the year of the first insider trading prosecution (Insider Trading Law) based on Bhattacharya and Daouk (2002) and for the year of the introduction of an electronic trading system (Electronic Trading System) based on Jain (2005), as well as (iv) an economic openness measure calculated as the ratio of imports plus exports over GDP (Trade Openness). Because equity market development is also likely influenced by the state and development of the local real economy, we also include as controls the annual growth in GDP per capita ($\triangle$GDP per Capita) and the GDP deflator ($\triangle$GDP Deflator) as a measure of inflation. To the extent that the different indicators needed updating, we did so by applying the same definitions as in the original sources to hand-collected data from the websites of the National Stock Exchanges. Appendix A describes the various measures in more detail. We expect all explanatory variables to have a positive impact on the MCAP/GDP and liquidity indicators, but a negative effect on the firm and industry concentration measures.

Panel A of Table 3.6 reports estimation results from a multivariate panel regression of each of the 5 yearly market development statistics on the different determinants and controls over the period 1990-2011, both without (left panel) and with (right panel) country fixed effects. While the model without fixed effects tells us which factors determine market development on average, the model with fixed effects captures factors that determine market development within a country over time. To control for possible dependence of residuals within a country, we cluster at the country level (see, e.g., Angrist and Pischke (2009); Cameron and Trivedi (2005)). However, when there are few clusters, in our case 16 country clusters, and few observations within a cluster, clustered standard errors might be biased (see, e.g., Wooldridge (2003), Cameron et al.

(2008), Cameron and Miller (2013)). So, we also briefly comment on what happens when we simply use heteroskedasticity adjusted standard errors.

When no fixed effects are included, the only measure that seems to systematically foster development is the dummy for the first prosecution of insider trading, leading to higher levels of liquidity (in terms percentage of non-zero returns) and lower sector concentration. The 'Political Constraints Index' and 'First Sign of Liberalization' dummy are significantly negatively related to sector concentration and the former negatively to the turnover ratio. Surprisingly, economic openness is associated with higher industry concentration. An explanation might be that in developing countries only few specific industries are engaged in international transactions and trade. When we instead focus on the results with fixed effects (RHS panel), i.e., on the variation within each country across time, we find the ratio of MCAP/GDP to be increasing with economic openness and a reduction in political risks (a higher political constraint index). Market liquidity improves with the EBRD transition index and following official liberalization (but actually decreases following first signs of liberalizations and with a higher political constraint index). Both sector and firm concentration decrease with the EBRD transition indicator and following first signs of liberalization. In unreported results, we also experiment with simple heteroskedasticity adjusted standard errors. The significance of most of the determinants improves and we additionally find the prosecution of insider trading and the implementation of an electronic trading system to positively affect MCAP/GDP and the former to also improve market liquidity.

Given our relatively short sample and many explanatory variables of which at least some are likely to be highly correlated, it should not come as a surprise that many market development determinants enter the regression insignificantly. To come to a more parsimonious model, we use model reduction techniques inspired by Hendry and Krolzig (2001)'s PCGets ("general-to-specific") methodology. For each market development indicator, we start by running a multivariate regression including all determinants of market development. Subsequently, we remove all insignificant regressors with a p-value above 10% if a joint F-test of the coefficients being equal to zero is insignificant at the 10% level. In case the test is significant, we only remove the regressor with the lowest absolute t-statistic and run a new model without this

regressor. We repeat all previous steps until we are left with only significant regressors.

Panel B of Table 3.6 reports estimation results for the reduced models, again both without (left panel) and with (right panel) country fixed effects. We report standard errors clustered at country level between brackets. Overall, we find that the market development indicators that were significant in the full model are also present in the reduced model. We do observe small changes in the estimates' magnitude and often increased significance. Rather than simply repeating the conclusions from the full model, we focus here on the economic magnitude of the estimates for the specifications with country-fixed effects. For most regressions, it is also the case that using the alternative standard errors produces virtually identical results. This is not surprising as country fixed effects may adequately account for within country correlation. As shown in the first column of the RHS of Panel B of Table 3.6, the MCAP/GDP is significantly positively related to economic openness, the political constraint index, the intensity of capital controls and first signs of liberalization. With respective standard deviations of 0.347 and 0.154, a one standard deviation increase in economic openness or a reduction in political risk (an increase in the political constraints index) separately lead to more than a 5.5 percentage points increase in MCAP/GDP (0.347×0.162 and 0.154×0.371). Similarly, following a one standard deviation increase in capital openness (1.448), we observe an increase of about 4.2 percentage points (1.448×0.029) in the MCAP/GDP and a 5.2 percentage points increase following a first sign of liberalization. Given an average MCAP/GDP of 16 percent, the effects are economically meaningful. The percentage of nonzero return is on average 80 percent but increases with about 8 percentage points (0.305×0.265) following a one-standard deviation increase in the EBRD transition indicator, and with 4.5 percentage points following an official liberalization. For firm and industry concentration, the EBRD transition indicator and the first sign of liberalization dummy have both a highly significant effect. A one standard deviation increase in the EBRD transition indicator (0.305) is associated with a 9 and 13 percentage points reduction in firm and industry concentration, respectively. Relative to mean values of 20 percent for firm concentration and 30 percent for industry concentration, these effects are highly economically significant. A first sign of liberalization has an even more pronounced economic effect on concentration measures, reducing firm and industry concentration by 17

and 25 percentage points, respectively. Finally and unexpectedly, the prosecution of insider trading increases (with 23.6 percent) rather than decreases firm concentration. In the case of turnover, we only find lower political risks to be associated with lower turnover. Overall, mostly the EBRD transition indicator and first signs of insider trading law seem to foster development within a country. Across countries, again, first sign of liberalization but also the prosecution of insider trading law are highly responsible for market development (see LHS of Panel B of Table 3.6).

3.4 The Diversification Benefits of CEE Equity Markets

In the early 90s, the emerging markets of primarily Latin America and South-East Asia were touted as the ideal investment, offering growth potential and great diversification benefits through the very low correlations they exhibited with developed markets (see, e.g., Harvey (1995)). Over the last two decades correlations between emerging markets and developed markets have greatly increased and their diversification benefits are much in doubt (see Bekaert and Harvey (2014), Christoffersen et al. (2012)). However, the CEE markets liberalized later and still showed relatively weak integration with world markets before the recent global financial crisis (see Bekaert et al. (2014)). Hence, they may still offer significant diversification potential. To examine this formally, we explore correlation dynamics, international hurdle rates and the importance of industry and country factors.

3.4.1 Correlation Dynamics

To consider correlation dynamics, we define 4 different benchmarks, the World market index, an overall European market index, the emerging market index and the Russian index, all from MSCI. The latter benchmark can reveal whether there is strong regional integration within Central and Eastern Europe. Table 3.7 tests for an upward trend in the quarterly non-overlapping return correlation, computed with weekly returns, of each of the CEE markets with the various benchmarks and confirms that correlations with respect to all benchmarks have substantially

increased over time.[76] In nearly all countries, we find a significant upward trend in correlations with benchmark index returns. Exceptions include Croatia and the Slovak Republic (with respect to all benchmarks), Romania and Ukraine (with respect to Emerging markets and Russia) and Slovenia (with respect to Russia only). A sub-sample analysis reveals that correlations mainly moved upwards from 2007 onwards. The increase is even higher for the aggregate CEE index than at the individual country level, growing from levels below 30% in the 97-02 sub-period to more than 80% (70%) with respect to global and European index returns. Correlations with respect to Emerging Markets and Russian index returns started at slightly higher levels (42% and 47%, respectively), and reach near unity (89% and 95%, respectively) towards the end of our sample. While the average full sample correlations are lower for frontier than for emerging markets, the upward trend is more pronounced for frontier markets.

It is well-known that correlations show substantial time variation and are sensitive to volatility changes, making them temporarily higher in any crisis (see, e.g., Ang and Bekaert (2002c)). Figure 3.1 plots a 4 quarter moving average of the benchmark quarterly correlations for the aggregate CEE index as well as for the median benchmark correlations across the individual CEE countries.

The patterns for the World, European and Emerging benchmarks are quite similar. While correlations inched up before 2003 already, correlations primarily showed large temporal swings. From 2003 onwards a clear upward trend is visible. Currently, correlations are over 80%, which is not too different from the correlation between the emerging market index and world market returns. If these correlation increases are largely permanent, then CEE markets have lost their diversification potential just as emerging markets more generally have. However, looking at the (unreported) top and bottom quartile of benchmark correlations across all CEE markets shows that the bottom quartile remains below 50% even towards the end of our sample, suggesting that considerable diversification benefits remain at the individual country level. As for the correlations with the Russian benchmark, only the median country correlations are of interest, because Russia is the main component of the CEE index. Correlations with Russia decreased

[76] In particular, we employ the linear time trend test by Bunzel and Vogelsang (2005) based on a simple time series model: $y_t = \beta_1 + \beta_2 t + u_t$, where y_t stands for the variable of interest and t for the linear time trend. We test for the null hypothesis of $\beta_2 = 0$ and use a Daniell kernel to estimate the standard errors in order to maximize the power of the test in small samples.

in the aftermath of the Russian crisis in 1998 and only started to trend up from 2005 onwards.

To help interpret the increase in correlations and to assess whether it is likely to be permanent or temporary, we decompose the correlation of a CEE market i with a benchmark b into three main components: the market's beta with respect to the benchmark ($\beta^b_{i,t}$), the benchmark's return volatility ($\sigma^2_{b,t}$), as well as country-specific ('idiosyncratic') volatility ($\sigma^2_{i,t}$) (see also Bekaert et al. (2009c)). Consider the following one factor model:

$$r_{i,t} = \alpha_i + \beta^b_{i,t} r_{b,t} + \varepsilon_{i,t}, \tag{3.4.1}$$

with $r_{i,t}$ and $r_{b,t}$ the returns on CEE market i and on the benchmark index, respectively, and $\varepsilon_{i,t}$ a country-specific shock. The model-implied correlation is then simply derived as:

$$\rho^b_{i,t} = \frac{\beta^b_{i,t}\sigma^2_{b,t}}{\sqrt{\left(\left(\beta^b_{i,t}\right)^2 \sigma^2_{b,t} + \sigma^2_{i,t}\right)\sigma^2_{b,t}}} = \frac{\beta^b_{i,t}\sigma_{b,t}}{\sqrt{\left(\left(\beta^b_{i,t}\right)^2 \sigma^2_{b,t} + \sigma^2_{i,t}\right)}} \tag{3.4.2}$$

This illustrates that the increase in correlations can result from (a combination of) increasing benchmark betas, an increase in the benchmark's volatility, or a reduction in country-specific risk. The high levels of volatility in developed markets since the start of the global financial crisis may therefore partly explain the high levels of correlations towards the end of the sample. Alternatively, further market development and integration may have pushed up market betas (see, e.g., Bekaert and Harvey (1997), Ng (2000), or Baele (2005b)) and, at the same time, reduced country-specific risk (see Baele and Inghelbrecht (2010)).

Table 3.7 therefore examines trends in quarterly betas and annualized idiosyncratic volatilities as well. Again, we measure both using weekly returns within the quarter. The betas of nearly all countries trend upwards, while idiosyncratic risk invariably tends to decrease over time. The benchmark betas of frontier markets are generally lower than those of emerging markets, but increase faster, as reflected by the higher average trend coefficient. Average country-specific risk is actually lower for frontier than for emerging markets, but trends more steeply downward, further contributing to a steeper increase in correlations of frontier relative

to emerging markets. The upward (downward) trend in benchmark betas (country-specific risk) suggests that CEE frontier markets are gradually becoming more integrated with world equity markets, a feature they do not share with many other frontier markets (see, e.g., Berger et al. (2011)).

The statistical significance in favor of an upward trend is stronger for the betas than for residual risk, but certainly not as strong as for the correlation trends. As Figure 3.2 shows, the upward trend is more pronounced for the betas of the CEE index than it is for the median country betas, which is to a large extent due to the higher upward trend in the betas of Russia and Turkey, the two largest contributors to the index. Figure 3.3 shows that the downward trend in residual volatility was interrupted on two occasions, namely during the Russian crisis and during the burst of the TMT bubble (1998-2001), and to a much smaller extent, during the global financial crisis (2008-2010).

We quantify the relative importance of the three different channels to explain the upward trend in market correlations in Table 3.8. We first report the 10-year (sample correlation) (ρ_{10}), the correlation over the last two years (ρ_2) and a proposed predicted correlation (ρ_{+1}; for one-year ahead correlation), based on the linear trend model of Table 3.7.[77] Note that the 2-year correlation is mostly but not always higher than the 10-year correlation, reflecting the general upward trend in correlations. We report the ρ_2/ρ_{10} ratio explicitly in column 4 of each block in Table 3.8. It is naturally mostly above 1.

In addition, we report the ratio of three model-based correlations over their 10-year sample correlation. The model-based correlation estimates are obtained by plugging alternate 2-year estimates of beta, benchmark volatility, and country-specific volatility in equation (3.4.2), while keeping the other two components at their 10-year value. Although the correlation increase likely comes from a simultaneous change in its three ingredients, we can thus visualize which component leads to the highest ρ_2/ρ_{10} ratio and therefore contributes the most to the increase. At the CEE index level, most of the increase in correlations over the last 10 years seems to be due to a decrease in region-specific volatility, and to a much smaller extent to an increase in

[77] In particular, we calculate the one-year ahead correlation using the estimated coefficients from the Bunzel and Vogelsang (2005) trend test estimated on quarterly data: $\hat{\rho}_t = \hat{\beta}_1 + \hat{\beta}_2 t$. Our one-year ahead forecast is then simply $\hat{\rho}_{+1} = \hat{\beta}_1 + \hat{\beta}_2(T+4)$, with T the number of quarterly observations at the end of the sample. Note that if the trend coefficient is not statistically significant we use the correlation for the whole sample period.

benchmark betas. Changes in benchmark volatility have either no effect (Europe) or even a slightly negative effect (Global, Emerging, Russia). The negligible effect of factor volatility is confirmed at the individual country level. For many of the smaller markets (Estonia, Hungary, Lithuania, Romania, Slovak Republic), however, increasing benchmark betas are a stronger source of increasing correlations than is decreasing idiosyncratic volatility.

We conclude that the increase in correlation does not appear due to a temporary increase in factor volatilities but rather due to declining idiosyncratic risk and increased betas. Such trends are more generally observed for emerging markets as well. Since it is therefore possible that correlations increase further, we use the trend model to estimate correlations one year ahead at the end of the sample period, and report them in the third column. Our model predicts correlations to be over 90% with the global and overall European market and the prediction for the correlation with the emerging market index is at 100% (that is, accounting for the trend moves the correlation above 1).

3.4.2 International Hurdle Rates

Table 3.7 also reports alphas for the different markets with respect to the various benchmarks. We find alphas to be always positive and economically large, but mostly statistically insignificant. The average alpha across all countries is 15.73% and similar in magnitude for emerging (16.09%) and frontier (15.37%) CEE markets. The high alphas could be either the result of a compensation for persistent country risk, or could simply reflect price appreciations following important historical improvements in the regulatory and trading environment. In the latter case, historical alphas likely overstate expected future alphas. That is why instead of looking at historical alphas, we prefer looking at hurdle rates, as they quantify diversification benefits in "expected return" space without actually using average historical returns.

A diversified mean-variance investor should invest in a new market as long as it improves the Sharpe ratio of her portfolio. Given a certain premium (expected excess return) on the portfolio the investor is holding, it is straightforward to compute the "hurdle" expected excess return on the foreign market the investor must exceed in order to increase the Sharpe ratio of her portfolio (see, e.g., Bekaert and Hodrick, 2011, Chapter 13). That is, given correlations and

volatilities, we can compute hurdle rates, defined as the lowest possible expected (excess) return for market i that must be earned for investors with a 100 percent investment in the benchmark to improve their Sharpe ratio when they invest in market i, given a specific expected (excess) return for the benchmark market. The hurdle rate HR_i^b (in excess of the risk-free rate) for country i with respect to benchmark b, is calculated as:

$$HR_i^b = \rho_{i,b}[E\left[r_b\right] - r_f^b]\frac{Vol\left(r_i\right)}{Vol\left(r_b\right)} = \beta_{ib}[E\left[r_b\right] - r_f^b], \tag{3.4.3}$$

with r_f^b, $E\left[r_b\right]$,and $Vol\left(r_b\right)$ the risk-free rate, expected return, and volatility of the benchmark market, respectively. $\rho_{i,b}$ is the correlation between returns on market i and those on the benchmark index, and $Vol\left(r_i\right)$ is local market volatility. Hurdle rates will be higher when the local market has a larger correlation with the benchmark, when the benchmark has a higher Sharpe ratio, and when the local market has a high total volatility. As a benchmark, we take MSCI World, assuming $[E[r_b] - r_f^b] = 5\%$. The key aspect of this computation is that it quantifies diversification benefits in "expected return" space without actually using average historical returns.

The first three columns of Table 3.9 report full sample correlations $\rho_{i,b}$, volatility ratios $(Vol\left(r_i\right)/Vol\left(r_b\right))$, and hurdle rates with respect to the world over the period 2002-2011. Subsequent columns show hurdle rates measured using data over the last 5 and 2 years as well as a 1-year forward looking hurdle rate.[78] Full sample hurdle rates are relatively low, and below 5 percent for most countries. The lowest values are observed in Latvia (2.08%), Croatia (2.50%), and Bulgaria (2.60%); the highest in Russia (7.13%), Hungary (6.42%), and Turkey (6.15%). The (more diversified) CEE index has a higher hurdle rate of 6.63%. Except for Latvia and the Slovak Republic, hurdle rates increase substantially when more recent data is used. The one-year ahead hurdle rate remains below 5% in 7 of the 13 countries, but rises to around 10% for Turkey and Russia, and to more than 7% for Hungary, Poland, and Romania. Given the large hurdle rates for Turkey and Russia, the two largest contributors to the CEE index, it is no surprise that the hurdle rate of the CEE index has increased further, to 8.59%.

[78]To calculate the forward-looking HR, we use the "beta representation" of the HR, namely $HR_i^b = \beta_{i,b}[E\left[r_b\right] - r_f^b]$, and obtain the one-year ahead HR by using the one-year ahead beta prediction from the trend model discussed in Table 3.7.

Therefore, to motivate investing in the CEE index or its large constituents, it is imperative that returns higher than in the overall market are earned. That is, CEE markets have become risky, high beta investments that should earn higher returns than standard benchmark indices.

Figure 3.4 plots the hurdle rate of the CEE index and of the median country from 1998 till 2011 and shows that hurdle rates for the median country are far below those of the CEE index. Figure 3.5 shows the 10-year hurdle rates geographically, with lighter colors indicating lower hurdle rates and better diversification opportunities. It is apparent that the more established markets have higher hurdle rates. We conclude that the diversification potential of mainly the larger and more developed markets has decreased substantially over time. Investors can, however, still reap diversification benefits from investing in some of the smaller markets.

3.4.3 Industry and Country Effects

Heston and Rouwenhorst (1994) use a simple decomposition of individual stock returns into world market, country and industry effects to illustrate the dominance of country effects in developed market returns. When firm variation is dominated by country factors, country diversification will be valuable. Various authors have, however, suggested that increasing country return correlations may have made industry effects more important, especially during the nineties (see, e.g., Cavaglia et al. (2000)). In contrast, Bekaert et al. (2009c) and Baele and Inghelbrecht (2009) suggest that the dominance of industry effects may have been temporary for developed markets and Phylaktis and Xia (2006b) show that country effects dominate for emerging markets. Yet, as the industry concentration results of Table 3.5 show, many of the CEE markets are dominated by a few industries, suggesting that industry effects may play a significant role in correlation dynamics. We therefore apply the analysis to the individual stocks returns of our CEE universe.

For each stock i that belongs to industry j and country k we define the following equation for each period t:

$$r_i = \alpha + \beta_1 I_{i1} + \beta_2 I_{i2} + ... + \beta_{10} I_{i10} + \gamma_1 C_{i1} + \gamma_2 C_{i2} + ... + \gamma_{14} C_{i14} + \varepsilon_i, \tag{3.4.4}$$

where industry dummy I_{ij} equals one if stock i belongs to industry j and zero otherwise, and

country dummy C_{ik} equals one if stock i belongs to country k and zero otherwise. We run into problems of multicollinearity if we estimate equation 3.4.4 directly by cross-sectional regressions for each week since each stock belongs both to a country and industry and industry and country dummies are defined for all 10 industries and 14 countries. In order to avoid having to choose one industry and country as a benchmark, Heston and Rouwenhorst (1994) propose to impose the following restrictions for each period t:

$$\sum_{j=1}^{10} w_j \beta_j = 0, \tag{3.4.5}$$

$$\sum_{k=1}^{14} v_k \gamma_k = 0, \tag{3.4.6}$$

where w_j and v_k denote the value weights of industry j and country k in the CEE market portfolio and $\sum_j w_j = \sum_k v_k = 1$. The estimated residuals of these OLS regressions are by construction orthogonal to all industry and country dummies. This implies that the average residual is zero in every industry and in every country. Since the CEE market index is simply the value-weighted average over all industries and countries, the estimated error term of the CEE index is also zero. By definition, the sum of all industries and countries for the CEE market index is zero such that the least-square estimate of α is equal to the value-weighted CEE market return.

We calculate weighted least squares (WLS) estimates for equation 3.4.4 each week subject to the restrictions in equations 3.4.5 and 3.4.6. The weekly cross-sectional regressions yield time series of the intercept, country and industry coefficients. The estimation procedure allows to decompose each value-weighted index of industry j (country k) into an effect common to all countries $\hat{\alpha}$, a pure industry (country) effect $\hat{\beta}_j$ ($\hat{\gamma}_k$), and the value-weighted average of country (industry) effects of the securities that make up its index:

$$r_j = \hat{\alpha} + \hat{\beta}_j + \sum_{k=1}^{14} \phi_{j,k} \hat{\gamma}_k C_{ik}, \tag{3.4.7}$$

$$r_k = \hat{\alpha} + \sum_{j=1}^{10} \theta_{j,k} \hat{\beta}_j I_{ij} + \hat{\gamma}_k, \tag{3.4.8}$$

where $\theta_{j,k}$ represents the proportion of total market capitalization of country k included in industry group j and $\phi_{j,k}$ represents the proportion of the capitalization of industry j pertaining to country k's stocks. In Table 3.10 and Figure 3.6, we investigate to what extent return variation at the individual stock level is mainly due to country or industry effects. Table 3.10 reports the annualized time-series volatilities of the value-weighted country (industry) index in excess of the value-weighted CEE index, the pure country (industry) effects and the value-weighted sum of industry (country) effects. The top panel shows that excess country volatility is high for all countries with an average of 35%. Most of the volatility of the value-weighted country indices can be attributed to country-specific effects. The annualized volatility of the pure country effects is on average 30.78% compared to the annualized volatility of the value-weighted sum of industry effects which is only 5.72%. The average annualized volatility of the excess value-weighted industry index (bottom panel) is lower than the one of the excess value-weighted country (21.5% compared to 35%). The volatility of pure industry effects is less than half the volatility of the pure country effects (14.9% versus 30.78%). Most of the volatility in industry indices can be attributed to the combined country effects than to pure industry effects (17.09% versus 14.90% on average).

The relative importance of country and industry effects may have changed over time. For instance, as markets become more developed, one would expect the relative importance of country effects to decrease, and the role of common CEE market and industry shocks to increase. Figure 3.6 plots a 1-year moving average of both the joint and individual explanatory power of the country and industry effects, as measured by the model's R^2s, for the country returns in excess of the common market effect. The pure country (industry) R^2 is calculated by setting the industry (country) coefficients to zero. The figure clearly shows that the importance of industry effects has increased over time. The total R^2 and the country R^2 seem to have dropped since the beginning of the millennium, while the industry R^2 has been constantly on the rise.

3.4.4 "International" versus "Non-International" stocks

A large literature suggests that international investors primarily focus on cross-listed, large and liquid stocks (see, i.e., Ammer et al. (2012) and Kang and Stulz (1997)). Miller (1999) and Foer-

ster and Karolyi (1999) show that announcements of cross-listings result in positive stock market reactions. According to the bonding hypothesis, cross-listed firms are valued higher because U.S. listings strengthen the legal protection of minority shareholders and reduce the agency costs of controlling shareholders (Stulz (1999), Coffee (2002), Reese and Weisbach (2002)). Controlling shareholders have thus an incentive to limit their private benefits, allowing firms to exploit valuable growth opportunities (Doidge et al. (2004)). At the same time, international cross-listings have been found to contribute to stock market development and integration (see, e.g., Errunza and Miller (2000), Bekaert et al. (2002), Karolyi (2004), and Edison and Warnock (2008)).

If international investors indeed predominantly focus on large, liquid, and cross-listed stocks, we would expect these stocks to have a higher correlation with respect to global benchmarks than small, illiquid "home" stocks. This "segmentation" may exist even in the larger and more developed markets. Similarly, large, liquid, cross-listed stocks listed in a small market may constitute an "integrated" submarket in an otherwise underdeveloped and segmented market. To test this hypothesis, we first divide our entire sample in (i) stocks with direct listings on other exchanges, ADRs or GDRs (cross-listed) and without cross-listings (home-only), (ii) stocks at the top 75th and bottom 25th quartiles of market capitalization, and (iii) stocks at the top 75th and bottom 25th quartiles of liquidity based on percentage of non-zero returns. Subsequently, we test whether (subcomponents of) correlations (Tables 3.7 and 3.8) and hurdle rates (Table 3.9) of small, illiquid, and home stocks with respect to global equity markets are indeed lower compared to those of large stocks, liquid stocks, and cross-listed stocks.

Table 3.7 reveals that cross-listed stocks have substantially higher market betas compared to home stocks (1.03 vs. 0.55 relative to the world), and generally steeper upward trends, suggesting that cross-listed stocks are indeed better integrated with global markets. The positive impact of betas on correlations is, however, largely eliminated by the cross-listed stocks' higher levels of idiosyncratic risk. Overall, we do not detect significant differences in benchmark correlations between home and cross-listed stocks. Interestingly, Table 3.8 shows that over the last 10 years correlations have increased faster for home than for cross-listed stocks, and that this is mainly due to a faster increase in benchmark betas. In other words, at least over the last

10 years, home stocks have been catching up with cross-listed stocks in their integration with global equity markets. While catching-up in betas also lead to a gradual convergence in hurdle rates (Table 3.9), we still find the hurdle rate based on the last 2 years of our sample period to be 1.45% higher for cross-listed stocks (5.88% vs. 4.43%).

The next rows of Tables 3.7, 3.8, and 3.9 show that large and liquid stocks have substantially higher benchmark correlations, betas, and hurdle rates. Trends in betas are significantly positive and large for large and liquid firms but much lower and mostly insignificant for small and illiquid firms. Similar to home stocks, the ratio of the last 2 to the last 10 year benchmark correlation (see Table 3.8) indicates that correlations have been increasing much faster for small than for large stocks. In fact, correlations measured over the final 2 years of our sample are roughly of the same magnitude for small and large stocks. While benchmark correlations of illiquid stocks have also benefitted from increasing benchmark betas, we still find benchmark correlations of the most liquid stocks to be nearly the double of those of the most illiquid stocks. Hurdle rates of small stocks, as measured over the final 2 years of our sample, have increased over time but are still below those of large stocks (5.88% vs. 4.43%). The persistently low benchmark correlations (betas) of the illiquid stocks imply hurdle rates below 3%, compared to more than 6% for the most liquid stocks.

3.5 The Cross-section of Expected Returns in CEE Markets

The literature on the cross-section of expected returns has mostly focused on the US and developed stock markets, yielding a number of well-known cross-sectional pricing effects that are inconsistent with the CAPM model. There is a lot less work in this area focusing on emerging and frontier markets. While early work by Claessens et al. (1995) did not detect size anomalies in emerging markets, Rouwenhorst (1999b), Barry et al. (2002b), and Cakici et al. (2013) show that emerging market returns share many cross-sectional anomalies with developed equity markets, in particular size, value, and momentum effects. The value effect appears more robust statistically and economically across these studies. For a sample of 24 frontier markets,

including 7 from the CEE area, de Groot et al. (2012) find evidence for a value and momentum but not for a size effect. In a sample of 11 CEE markets, Zaremba and Konieczka (2014) find strong value and size effects but no momentum effect. This is confirmed by Cakici et al. (2013) who, studying 18 emerging markets, find that the momentum effect does not extend to Eastern Europe.

In this section, we examine the size, value, momentum, volatility, betting against beta and liquidity effects for our sample of CEE stocks. The methodology and results are described in Section 3.5.1. The cross-sectional strategies may lead to portfolios that significantly outperform market indices. We consider this formally in Section 3.5.2, applying the Brandt et al. (2009) parametric portfolio policy methodology, with the firms' size, book-to-market, past performance, realized volatility, local and global market beta, and liquidity as characteristics.

3.5.1 Cross-sectional Pricing Effects

To the extent possible, we follow the methodologies in the original papers documenting the effects, but, given the limited number of stocks, it is impossible to stratify stocks in deciles or even quintiles. Therefore, we sort stocks based on the various characteristics in quartiles and compare the performance of the top and bottom quartiles, value-weighting within the quartiles. Appendix B describes the procedure for creating the portfolios in more detail. To correct for risk, we consider regressions of the form:

$$r_{Pt} - r_{ft} = \alpha_P + \beta_P \left(r_{Mt} - r_{ft}\right) + \text{additional controls } + \varepsilon_{Pt}, \tag{3.5.1}$$

where r_{Pt} and r_{ft} are the monthly portfolio and risk-free return (in USD) in month t, r_{Mt} the return on a benchmark market portfolio, β_P the corresponding slope coefficient, and ε_{Pt} a white noise error term. The regression's intercept, α_P, is a measure of risk-adjusted performance. As market factor, we use returns on indices proxying for respectively global, European, CEE, and Emerging markets. We also run two specifications with either the CEE and European market factor or the global size, value, and momentum factors (see, e.g., Fama and French (1998)) as controls in addition to the global market factor. We report Newey-West standard errors.[79]

[79] We use $0.75T^{1/3}$ as the truncation lag, with T the total number of observations.

Our exercise resolutely takes the perspective of a global investor investigating the pricing of CEE stocks as a whole, relative to global, regional and emerging market wide indices. The markets are simply too small to consider ranking stocks within countries.[80] Of course, given evidence that some of our markets are still partially segmented (see Bekaert et al. (2014)), any alphas we find do not indicate the existence of pricing anomalies within a particular CEE country.

Panel A of Table 3.11 reports results when no restrictions are imposed on the portfolios with respect to country/industry concentration, while in Panels B and C we impose the different portfolios to have either the same industry or country weights as the aggregate CEE index. For each strategy, we report the average return and volatility of the high, low and spread portfolios. We also report alphas with respect to various (combinations of) risk benchmarks. In addition, we record the correlation of the various portfolios with the world market, its volatility ratio relative to the world market's volatility and the hurdle rate for a world investor, assuming a 5% premium on the world market. The final column characterizes the liquidity of the portfolio by providing the average proportion of non-zero returns for stocks in each portfolio. This proportion is measured monthly but rebalancing takes place quarterly. More details on how each cross-sectional portfolio is constructed are relegated to Appendix B.

3.5.1.1 Size Effect

Banz (1981) and Reinganum (1981) were the first to document that stocks with lower market capitalization (small stocks) tend to have higher average returns. In their influential article, Fama and French (1992) found that the smallest size decile of US stocks outperformed the largest by 0.74% per month. While several studies, including Hirshleifer (2001) and Schwert (2003), argued that the effect disappeared in the US soon after its discovery, there is substantial support for a size effect in non-US developed and emerging markets, even today (see van Dijk (2011) for an overview).

The first row of Panel A of Table 3.11 shows that the average return on a value-weighted portfolio containing the 25% smallest stocks equals 24.6%, which is 8.7% higher than the

[80] Lischewski and Voronkova (2012) find evidence for value and size, but not for liquidity effects in Poland, one of the largest CEE markets.

return on a portfolio containing the 25% largest stocks. Annualized volatility is higher for the large cap than for the small cap portfolio (38.6% vs 32.7%). While long-only portfolios in the bottom and top size quartiles both generate positive alphas for various (combinations of) risk controls, they are, apart from being much larger in magnitude, only significant for the small stock portfolio. The alphas of the spread portfolio are economically large but significant only when controlled for by the CEE or CEE/EU/World market benchmarks. Because of its lower volatility and correlation with global market returns, the small stock portfolio has a substantially lower hurdle rate relative to the large cap portfolio (4.77% vs 7.28%). The last column shows that large caps are on average more liquid than small stocks, as reflected by their higher percentage of non-zero returns (84.4% vs 64.3%). Imposing the same industry exposure on the different portfolios as the CEE index leads to an increased performance of the large cap portfolio, and a further decrease in alphas for the spread portfolio. Imposing the same country structure, however, leads to a proportionally larger increase in the performance of small relative to large caps, so that the spread portfolio now has significant alphas of more than 13%, annually, in three cases.

We conclude that CEE size portfolios generate mostly significant alphas because of segmentation; the risk factors do not capture their returns, and this effect is overall somewhat stronger for small cap portfolios. The spread portfolio does generate high alphas, but these alphas are not generally statistically significant.

3.5.1.2 Value Effect

The value effect refers to the observation that value stocks, that is, stocks with high ratios of a fundamental like book value or cash flow to price, have higher average returns than growth stocks, which have low ratios of fundamentals to price (Bondt and Thaler (1985), Fama and French (1992) and Lakonishok et al. (1994)). Recent work by Asness et al. (2013), amongst many others, shows that the value effect continues to exist, both in the US and in other (developed) markets.

Rows 2 to 4 of Panel A of Table 3.11 show that portfolios containing the 25% stocks with the highest dividend yield (DY), or the lowest price earnings (P/E) or market-to-book (M/B)

ratios, i.e. value stocks, significantly outperform portfolios containing stocks with the 25% lowest DY, or highest P/E or M/B ratios ('growth stocks').[81] The difference is substantially larger when portfolios are sorted based on P/E or M/B ratios (alphas between 20% and 25%) than on DY. Hurdle rates are similar for value and growth stocks (about 7%), and much lower than the average return on the value portfolio. We obtain similar results for the country or industry neutral strategies based on the P/E and M/B ratio, except that low dividend yield stocks now perform better. The alphas on the spread portfolios are mostly insignificant for the dividend yield sorted portfolios, and almost always significant for the P/E ratio sorts. For the market to book sorted portfolios, they are always statistically significant when the industry or country structure of the CEE index is matched. The differential performance between value and growth stocks does not seem to be driven by differences in liquidity, as both have high percentage of non-zero return ratios.

3.5.1.3 Momentum

Cross-sectional momentum was first documented by Jegadeesh and Titman (1993), who showed that stocks that performed well over the previous 3 to 12 months ('winner stocks') tend to outperform stocks that performed poorly ('loser stocks). Rouwenhorst (1999b) and more recently Asness et al. (2013) provide strong evidence for momentum in international markets.

We do, however, not find any evidence for momentum in CEE equity markets, irrespective of whether we sort based on the stocks' performance in months $t-12$ to $t-1$ or $t-6$ to $t-1$ (columns 5-6 in Table 3.11). In fact, we find generally past losers to outperform past winners. The effect is even statistically significant for one year momentum returns except when the CEE's country structure is imposed. For the 6 month momentum portfolios, the results vary a lot across weighting schemes. Our results are in line with those in Cakici et al. (2013) and Zaremba and Konieczka (2014), who also did not find momentum effects in CEE equity markets.

[81] We lag all characteristics by 1 quarter to ensure they were in the information set at the time of rebalancing.

3.5.1.4 Low Volatility and Low Beta Effect

One of the strongest empirical findings is that low volatility and low beta stocks tend to outperform high volatility and beta stocks. Already in the 1970s, research showed that the relationship between risk as measured by market beta and return was much flatter than predicted by the CAPM (Black (1972) and Black et al. (1972)), or even downward sloping (Haugen and Heins (1975)). Frazzini and Pedersen (2014) show that this effect is also present in recent data and across different markets and asset classes. Ang, Hodrick, Xing, and Zhang (2006a; 2009b) reveal that stocks with recent past high idiosyncratic volatility underperform stocks with low past idiosyncratic volatility, both in the US and international markets. Frazzini and Pedersen (2014) show that this is also the case for low relative to high beta stocks. Blitz and Vliet (2007) and Blitz et al. (2013) confirm these findings for developed and emerging markets, respectively, but de Groot et al. (2012) do not find low volatility or low beta effects in their sample of frontier markets.

Table 3.11 (row 7) offers only limited evidence for a volatility effect in CEE markets. Low volatility stocks, where volatility is measured over the previous quarter, outperform high volatility stocks in the unrestricted and country-weights case, but the outperformance is not statistically significant. We do not find that stocks with low betas (either measured with respect to the local (row 8) or global (row 9) equity market, using annual windows and weekly returns outperform high beta stocks, or vice versa.

3.5.1.5 Liquidity

In Section 3.3.1, we showed that market liquidity, as measured by market turnover and the proportion of non-zero trading days, varies widely across CEE markets, and that it is generally below the levels observed in developed markets. Following Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), Datar et al. (1998a), Chordia et al. (2001), and Liu (2006), a vast literature has shown that liquidity is a priced risk factor. In a sample of 19 emerging markets (mostly from Latin-America and South-East Asia), Bekaert et al. (2007b) find that unexpected liquidity shocks are positively correlated with contemporaneous stock returns and negatively with dividend yields, suggesting that liquidity is a priced risk factor also

in emerging markets.

The final line of Table 3.11 reports performance statistics for portfolios containing stocks belonging to the top and bottom liquidity quartiles. As a liquidity measure, we use the proportion of non-zero returns. Even though we eliminated the most illiquid stocks from our sample, we observe a large difference in liquidity between the top and bottom quartile. Stocks in the most liquid bucket have non-zero returns on 90% of trading days, compared to 'only' 46% for the stocks belonging to the bottom bucket. The alphas of the restricted low liquidity portfolios are typically more than 20%, about 5 to 10% higher than those of the high liquidity portfolio. In the unrestricted case, the alphas of the spread portfolio range from 5.3% to 7.2%, and are never significant. When the CEE's index industry structure is imposed, the alphas of the spread portfolio almost double and become highly statistically significant.

3.5.1.6 Free-Float

As discussed in Section 3.2.4, government agencies still play a relatively large role in CEE equity markets, and free-float market weights differ substantially from total market weights. Table 3.12 compares the performance of indices that are based on the full sample (Original) and on a sample of stocks for which free-float data is available, either using (rescaled) market weights as before (FF1) or using free-float market weights (FF2). Because free-float data is not available before, the sample starts in 2003 only (rather than in 1996). We find that the size effect becomes weaker when we focus on free-float data and loses its significance when we use free-float market weights. This effect might be due to the fact that we mostly select larger firms for the free-float sample. The value effect remains robust when we sort based on the market to book ratio. We still find a reverse mostly insignificant momentum effect. The low volatility effect is a bit stronger in this shorter sample than it is in the larger sample underlying the results in Table 3.11; it generates large and significant alphas for the spread portfolios relative to the CEE/EU/World benchmark but not relative to the global Fama and French factor benchmark. The volatility effect is somewhat weaker but nonetheless preserved in the free-float adjusted samples. Free-float adjustments do not change the conclusion that we find no statistical evidence for a low beta effect in CEE stocks. The illiquidity effect is also very

similar across samples, only leading to slightly lower returns on the illiquid portfolios. In sum, although free-float adjustments seem to change the results slightly, our main conclusions are not affected.

3.5.2 A CEE portfolio with Cross-sectional Tilts

Because the CEE markets have high correlations with the world market, they require rather high hurdle rates before a world investor would want to include them in her portfolio. If we run a regression of our CEE index return on the world market return over the full sample from 1996 to mid 2011, we find a beta of 1.24 and an annualized alpha of 11.60 which is, however, not statistically significant (t-stat of 1.178). But section 3.5.1 revealed that some stock market characteristics may be associated with higher returns. In this section, we create a CEE portfolio that exploits some of these pricing anomalies and may perform better than the index. Our methodology is based on the work of Brandt et al. (2009). They maximize a statistical function of the portfolio return, where the portfolio weights are modeled as benchmark weights (e.g. value-weights) plus an active component that over- or underweights stocks based on their characteristics (such as size, value,...). This methodology only requires finding the coefficients on the characteristics that maximize the in-sample utility, making an inherently complex optimization problem feasible. In our application, we modify the Brandt et al. (2009) framework to a tracking error framework, with the CEE index as the benchmark.

To understand our approach, we need to establish some notation. At each point in time t, we need to find the optimal weights $w_{i,t}$ for a portfolio of N_t stocks. The portfolio's return is simply calculated as $r_{p,t+1} = \sum_{i=1}^{N} w_{i,t} r_{i,t+1}$, with $r_{i,t+1}$ being the return on stock i observed over the period t to $t+1$. The return on the benchmark portfolio is calculated as $r_{b,t+1} = \sum_{i=1}^{N} w_{i,t}^{b} r_{i,t+1}$, with $w_{i,t}^{b}$ being the stock i's benchmark weight at time t. Similar to Brandt et al. (2009), we model the portfolio weights as the sum of the benchmark weight and an active component that over- or underweights a stock based on its characteristics $x_{i,t}$:

$$w_{i,t} = w_{i,t}^{b} + \frac{1}{N_t} \theta^T x_{i,t} \tag{3.5.2}$$

with θ being a vector of coefficients that translates the characteristics in over- or underweights. Because the characteristics $x_{i,t}$ are standardized cross-sectionally to have a zero mean and unit standard deviation across all stocks at date t, the deviations from the benchmark always sum up to zero, and the total weights to 100%. By normalizing $\theta^T x_{i,t}$ by N_t, the total active bet does not increase with the number of stocks.

Our investor maximizes the following objective function with respect to θ:

$$E\left[r_{p,t+1} - r_{b,t+1}\right] - \frac{\gamma}{2} Var\left[r_{p,t+1} - r_{b,t+1}\right] \tag{3.5.3}$$

While equation (3.5.3) looks like a mean-variance utility function, the return is the alpha over the benchmark portfolio and the risk is measured relative to the benchmark. In other words, this is a standard tracking error problem that portfolio managers would solve seeking to outperform the CEE benchmark.

As shown in Appendix C, taking first-order conditions and using sample moments to estimate the relevant theoretical moments yields the optimal θ:

$$\theta = \frac{1}{\gamma}\left[\frac{1}{T}\sum_{t=0}^{T-1}(x_t R_{t+1} - \overline{xR})(x_t R_{t+1} - \overline{xR})'\right]^{-1}\overline{xR} \tag{3.5.4}$$

where x_t is a k by N_t matrix containing the k (cross-sectionally standardized) characteristics for the N_t stocks observed at time t, R_t the corresponding $(N_t \times 1)$ vector of returns and $\overline{xR} = \frac{1}{T}\sum_{t=0}^{T-1} x_t R_{t+1}$ the average cross-product between characteristics and returns. Individual θ's will be positive (negative) to the extent that the corresponding characteristics correlate positively (negatively) with returns. The coefficient γ governs the portfolio manager's flexibility to deviate from the benchmark: lower (higher) values for γ will be associated with more (less) extreme tilts away from the benchmark. In our empirical implementation, we set $\gamma = 5$.

Because short selling in CEE markets is likely to be costly and only feasible for the most

liquid stocks[82], we only present results for a policy than truncates portfolio weights at 0:

$$w_{i,t}^{+} = \frac{\max\left[0, w_{i,t}\right]}{\sum_{j=1}^{N_t} \max\left[0, w_{j,t}\right]} \tag{3.5.5}$$

where we divide by the sum of the positive weights to make the portfolio weights sum up to 100 percent. We truncate the weights obtained from the unconstrained optimization ex-post; Embedding the short-sale constraints directly within the optimization only improves performance marginally but renders our bootstrapping exercise (see below) unstable.

For the analysis, we use monthly stock returns as well as firm-level characteristics for 2,090 stocks between January 1996 and May 2011. Note that we use all available stocks in our analysis and do not restrict the sample to the stocks available in our tailor-made CEE index, which represent 85% of total market capitalization. For each firm, we construct the following firm characteristics at the end of each month: the log of market capitalization of the previous quarter (*me*), the log of one plus the market to book ratio of the previous quarter (*mtbv*), the price return in US dollar between months $t-12$ and $t-1$ (*mom*), the volatility of daily returns over the previous quarter (*vol*), the percentage of zero daily (price) returns in local currency for each stocks over the previous year (*illiq*), and the beta from a regression of weekly returns on the value-weighted CEE index (local) over the previous year (*beta*). Before standardizing characteristics, we winsorize each characteristic at the 1% and 99% level across all stocks at each point in time in order to account for possible outliers. The benchmark weights correspond to the weight that each stock has at each point in time in a value-weighted index of all 2,090 stocks (to the extent that the stock is part of the sample at that specific point in time). We calculate bootstrapped standard errors using the procedure described in Brandt et al. (2009).[83]

We present estimation results and performance statistics in Panel A and B of Table 3.13, respectively. The second row of Panel A of Table 3.13 reports parameter estimates and port-

[82]Daniel and Lhabitant (2012) reveal that the ability to short-sell in CEE markets is limited and mostly executed by hedge funds. While short-selling is possible to some extent in the larger Eastern European countries such as Turkey, Poland, Hungary, Russia, and the Czech Republic (ordered by the ease of short-selling), it is very limited in smaller and less developed markets such as the Slovak Republic, Latvia, Estonia, Romania, and Slovenia.

[83]In particular, we generate a large number of samples of returns and characteristics. We take random draws (with replacement) from monthly observations for each stock returns and its characteristics using only liquid and selected observations of each stock. For each sample, we estimate the coefficients of the optimal portfolio and compute the covariance matrix of the coefficients across all the bootstrapped samples (1000 replications).

folio characteristics for a parametric portfolio policy that includes size, value, and momentum characteristics. Columns 1 to 3 report optimal estimates for θ as well as bootstrapped standard errors. For the size and market-to-book characteristic, we find negative θ's, meaning that the portfolio policy overweights small and value firms but underweights large and growth firms. The θ for momentum is positive, tilting the portfolio towards past winners but away from past losers. Because the firm characteristics are standardized across firms, the relative importance of the characteristics can be read from the coefficients' absolute sizes. We find the size effect to dominate the value effect, which in turn dominates the momentum effect. While bootstrapped standard errors are rather large, we find the percentage of simulated thetas that are of the opposite sign to be 2.5% for the value, 8.3% for the size, and 12% for the momentum characteristic. The next columns provide more insights in the optimal portfolio weights and characteristics. Compared to the value-weighted CEE benchmark portfolio (row 1), the optimal portfolio has a lower average maximum portfolio weight (10.2% versus 17%). To get a better sense of the composition of the optimized portfolio, the last three columns report the average of the weighted characteristics of the portfolio calculated as $\sum_{i=1}^{N_t} w_{i,t}^{+}\hat{x}_{i,t}$. The benchmark portfolio has a large bias towards large firms (because of value-weighting) and positive but smaller tilts towards firms with high market-to-book ratios (growth firms) and past winners. In contrast, the optimized portfolio has a bias towards small firms, value firms, and past winners.

Panel B of Table 3.13 reports performance statistics for the 3-characteristics optimal parametric portfolio policy. The optimal portfolio has both a substantially higher annualized mean excess return (23.3% versus 15.3%) and a lower annualized volatility (32.5% versus 37.1%) relative to the benchmark market-weighted CEE benchmark index. This leads to a Sharpe Ratio that is 0.31 percentage points higher for the optimal compared to the benchmark portfolio (0.72 versus 0.41). The next columns report annualized alphas from univariate regressions of portfolio returns on our tailor-made CEE index (85% market capitalization threshold), the MSCI EM index, the MSCI Europe index, the MSCI World index, as well as multivariate regressions on the latter three indices and on the global Fama-French factors and momentum. The first row shows that even the simple value-weighed CEE index has relatively large (but insignificant)

alphas with respect to all benchmarks.[84] For instance, an investment in the CEE index yields an alpha of 11.3%, correcting for the global Fama-French factors and momentum. We interpret this as a sign of segmentation, as global risk factors largely fail to capture CEE equity market returns. The second row shows that the parametric portfolio policy nevertheless generates substantially larger alphas than the simple value-weighted CEE index. Optimally taking into account size, value, and momentum characteristics leads to an outperformance of 13.9% relative to the benchmark CEE index. The strategy's alpha after correcting for the global Fama-French and Momentum risk factors equals 19.3%, which is nearly 8% points higher than the benchmark CEE index' alpha. The information ratio (relative to the benchmark CEE model) increases from 0.65 for the value-weighted CEE index to 1.86 for the parametric portfolio policy, a very substantial increase. The hurdle rate of the optimized portfolios with respect to the global equity market is lower than for the benchmark CEE index (6.73% versus 7.24%), mainly because of its lower volatility.

The third row of Panel A of Table 3.13 reports parameter estimates and portfolio characteristics for a parametric portfolio policy that includes 3 additional factors, namely lagged one-year volatility, lagged one-year liquidity, and the local beta with respect to our tailor-made CEE index. The policy tilts the portfolio towards low volatility stocks, capturing the low volatility effect, and towards less liquid stocks, potentially capturing an illiquidity risk premium. Contrary to what is typically found in developed markets, our strategy tilts weights towards high not low beta stocks.[85] We continue to find that the optimal portfolio is tilted towards small stocks, value stocks, and past winners. Now not only the θ for value but also for momentum is statistically significant at the 5% and 10% level, respectively, while size, illiquidity, and local beta are borderline significant at the 10% level.[86] Figure 3.7, which plots the value-weighted portfolio characteristics over time, shows that the reported tilts are relatively stable over time, and as such, that the strategy is making consistent bets over time.

[84]Note that the value-weighted CEE index, which involves all stocks, has a positive but small (3.75%) alpha with respect to the CEE benchmark, which only contains stocks (from large to low) until a 85% of total market capitalization is reached.

[85]We continue to find this even when we eliminate the likely correlated volatility characteristic. In fact, it is even true in a univariate exercise focusing on beta.

[86]When we examine the characteristics univariately, market to book, volatility, beta, and illiquidity, all yield statistically significant θ 's; the θ for momentum is borderline significant, but the one for size is not.

Panel B of Table 3.13 shows that adding volatility, illiquidity, and beta characteristics improves the performance of the strategy only marginally. Relative to the 3-characteristics case, the mean return of the strategy based on 6 characteristics increases from 23.3% to 25.3%. However, because also the strategy's volatility increases from 32.5% to 34.6%, the Sharpe Ratio increases with 1 percentage point only from 0.72 to 0.73. The alpha with respect to CEE benchmark increases with 1.4% points (from 13.85 to 15.27%); the alpha with respect to the global Fama-French and momentum model with 3.03% (from 19.26% to 22.28%). Because of the strategy's larger volatility, the information ratio (with respect to the CEE benchmark) only increases slightly (from 0.1.86 to 1.91) and the hurdle rate increases from 6.71% to 7.05%.

3.6 Conclusions

This paper provides a comprehensive and detailed analysis of Central and Eastern European (CEE) equity markets from the mid-1990s until now. We use firm-level data to create custom-made indices and indicators to maximize coverage. We find that there are substantial differences across different CEE indices. While for the overall index, the correlation between index returns across data vendors and our index returns is well over 90%, it can be below 50% for some of the smaller markets. We show that there is considerable heterogeneity in the degree, dynamics, and determinants of market development across the different markets. Using market size, liquidity and concentration indices, Russia, Turkey, Hungary and Poland are the most developed markets; Serbia, Latvia, the Slovak Republic, and Kazakhstan are least developed. One institutional feature that contributes most robustly to differential market development across countries is the implementation of insider trading laws.

Studying the diversification benefits to global investors, we find that CEE markets have experienced similar trends as emerging markets more generally, with strongly increasing correlations with global benchmarks over time. Increased correlations can occur because of higher benchmark volatilities, higher betas, or lower country-specific risks. We find that changes in correlations are primarily driven by increasing betas and reduced idiosyncratic risks, with their relative contributions varying across countries. For smaller markets such as Estonia, Hungary, Lithuania, Romania and the Slovak Republic, increasing betas dominate, but these markets still

feature relatively low correlations with global benchmarks. At the CEE index level, reduced idiosyncratic risk plays a larger role and correlations have increased to over 80%. Finally, we examine the pricing of various stock specific characteristics and show that there are substantial premiums associated with investing in small, value, low volatility and illiquid stocks. We show how an active strategy can tilt a CEE investment towards these stock characteristics, earning an annualized alpha of 15.3% with respect to the CEE benchmark and of 22.3% with respect to the global Fama-French and momentum model. The information ratio of this strategy is more than double that of the passive value-weighted portfolio.

References

Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Ammer, John, Sara B Holland, David C Smith, and Francis E Warnock, 2012, Us international equity investment, *Journal of Accounting Research* 50, 1109–1139.

Ang, Andrew, and Geert Bekaert, 2002c, International asset allocation with regime shifts, *Review of Financial Studies* 15, 1137–1187.

Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006a, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2009b, High idiosyncratic volatility and low returns: International and further U.S. evidence, *Journal of Financial Economics* 91, 1–23.

Angrist, Joshua D, and Jörn-Steffen Pischke, 2009, *Mostly harmless econometrics: An empiricist's companion* (Princeton University Press).

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Baele, Lieven, 2005b, Volatility spillover effects in European equity markets, *Journal of Financial and Quantitative Analysis* 40, 373–402.

Baele, Lieven, and Koen Inghelbrecht, 2009, Time-varying integration and international diversification strategies, *Journal of Empirical Finance* 16, 368–387.

Baele, Lieven, and Koen Inghelbrecht, 2010, Time-varying integration, interdependence and contagion, *Journal of International Money and Finance* 29, 791–818.

Banz, Rolf W., 1981, The relationship between return and market value of common stock, *Journal of Financial Economics* 9, 3–18.

Barry, Christopher B., Elizabeth Goldreyer, Larry Lockwood, and Mauricio Rodriguez, 2002b, Robustness of size and value effects in emerging equity markets, 1985-2000, *Emerging Markets Review* 3, 1 – 30.

Bekaert, Geert, Michael Ehrmann, Marcel Fratzscher, and Arnaud Mehl, 2014, The global crisis and equity market contagion, *Journal of Finance* 69, 2597–2649.

Bekaert, Geert, and Campbell R. Harvey, 1997, Emerging equity market volatility, *Journal of Financial Economics* 43, 29–77.

Bekaert, Geert, and Campbell R Harvey, 2004, A chronology of important financial, economic and political events in emerging markets., available at *http : //people.duke.edu/ charvey/Country_risk/chronology/chronology_index.htm*.

Bekaert, Geert, and Campbell R Harvey, 2014, Emerging equity markets in a globalizing world, available at SSRN: http://ssrn.com/abstract=2344817.

Bekaert, Geert, Campbell R Harvey, and Robin L Lumsdaine, 2002, Dating the integration of world equity markets, *Journal of Financial Economics* 65, 203–247.

Bekaert, Geert, Campbell R Harvey, and Christian Lundblad, 2005b, Does financial liberalization spur growth?, *Journal of Financial Economics* 77, 3–55.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad, 2007b, Liquidity and expected returns: Lessons from emerging markets, *Review of Financial Studies* 20, 1783–1831.

Bekaert, Geert, and Robert J. Hodrick, 2011, *International financial management* (Pearson Prentice Hall), (2nd edition) edition.

Bekaert, Geert, Robert J. Hodrick, and Xiaoyan Zhang, 2009c, International stock return comovements, *Journal of Finance* 64, 2591–2626.

Berger, Dave, Kuntara Pukthuanthong, and Jimmy Yang, 2011, International diversification with frontier markets, *Journal of Financial Economics* 101, 227–242.

Bhattacharya, Utpal, and Hazem Daouk, 2002, The world price of insider trading, *Journal of Finance* 57, 75–108.

Black, Fisher, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444–454.

Black, Fisher, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, Studies in the Theory of Capital Markets. Michael C. Jensen, ed. New York, Praeger, pp. 79-121.

Blitz, David, Juan Pang, and Pim Van Vliet, 2013, The volatility effect in emerging markets, *Emerging Markets Review* 16, 31–45.

Blitz, David, and Pim Van Vliet, 2007, The volatility effect: Lower risk without lower return, *Journal of Portfolio Management* 34, 102–113.

Bondt, Werner F.M. De, and Richard H. Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

Brandt, Michael W, Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447.

Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441–464.

Bunzel, Helle, and Timothy J. Vogelsang, 2005, Powerful trend function tests that are robust to strong serial correlation, with an application to the Prebisch-Singer hypothesis, *Journal of Business & Economic Statistics* 23, 381–394.

Cakici, Nusret, Frank J. Fabozzi, and Sinan Tan, 2013, Size, value, and momentum in emerging market stock returns, *Emerging Markets Review* 16, 46 – 65.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller, 2008, Bootstrap-based improvements for inference with clustered errors, *Review of Economics and Statistics* 90, 414–427.

Cameron, A Colin, and Douglas L Miller, 2013, A practitioner's guide to cluster-robust inference, *forthcoming in Journal of Human Resources* 221–236.

Cameron, A Colin, and Pravin K Trivedi, 2005, *Microeconometrics: methods and applications* (Cambridge University Press).

Campbell, John Y, Andrew Wen-Chuan Lo, and Archie Craig MacKinlay, 1997, *The econometrics of financial markets* (Princeton University Press).

Caporale, Guglielmo Maria, and Nicola Spagnolo, 2011, Stock market integration between three CEECs, Russia, and the UK, *Review of International Economics* 19, 158–169.

Cavaglia, Stefano, Christopher Brightman, and Michael Aked, 2000, The increasing importance of industry factors, *Financial Analyst Journal* 56, 41–54.

Chelley-Steeley, Patricia L., 2005, Modeling equity market integration using smooth transition analysis: A study of Eastern European stock markets, *Journal of International Money and Finance* 24, 818–831.

Chinn, Menzie D, and Hiro Ito, 2006, What matters for financial development? Capital controls, institutions, and interactions, *Journal of Development Economics* 81, 163–192.

Chinn, Menzie D, and Hiro Ito, 2008, A new measure of financial openness, *Journal of Comparative Policy Analysis* 10, 309–322.

Chordia, Tarun, Avanidhar Subrahmanyam, and Ravi Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3–32.

Christoffersen, Peter, Vihang Errunza, Kris Jacobs, and Hugues Langlois, 2012, Is the potential for international diversification disappearing? A dynamic copula approach, *Review of Financial Studies* 25, 3711–3751.

Claessens, Stijn, Susmita Dasgupta, and Jack Glen, 1995, Return behavior in emerging stock markets, *World Bank Economic Review* 9, 131–151.

Coffee, John C. Jr, 2002, Racing towards the top?: The impact of cross-listings and stock market competition on international corporate governance, *Columbia Law Review* 1757–1831.

Daniel, Jeannine, and François-Serge Lhabitant, 2012, *Handbook of Short Selling*, chapter 11: Reflections on Short Selling Regulations in Western and Eastern Europe (Elsevier).

Datar, V., N. Naik, and R. Radcliffe, 1998a, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203–219.

Datastream, 2008, *Datastream Global Equity Inidices, User Guide*, Thomson Reuters.

de Groot, Wilma, Juan Pang, and Laurens Swinkels, 2012, The cross-section of stock returns in frontier emerging markets, *Journal of Empirical Finance* 19, 796–818.

Doidge, Craig, G Andrew Karolyi, and René M Stulz, 2004, Why are foreign firms listed in the US worth more?, *Journal of Financial Economics* 71, 205–238.

Edison, Hali J, and Francis E Warnock, 2008, Cross-border listings, capital controls, and equity flows to emerging markets, *Journal of International Money and Finance* 27, 1013–1027.

Errunza, Vihang R, and Darius P Miller, 2000, Market segmentation and the cost of the capital in international equity markets, *Journal of Financial and Quantitative analysis* 35, 577–600.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1998, Value versus growth: The international evidence, *Journal of Finance* 53, 1975–1999.

Foerster, Stephen R, and G Andrew Karolyi, 1999, The effects of market segmentation and investor recognition on asset prices: Evidence from foreign stocks listing in the united states, *Journal of Finance* 54, 981–1013.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Gilmore, Claire G., Brian M. Lucey, and Ginette M. McManus, 2008, The dynamics of Central European equity market comovements, *The Quarterly Review of Economics and Finance* 48, 605–622.

Gjika, Dritan, and Roman Horvath, 2012, Stock market comovements in Central Europe: Evidence from asymmetric DCC model, *William Davidson Institute Working Paper No.1035* .

Harvey, Campbell R, 1995, Predictable risk and returns in emerging markets, *Review of Financial Studies* 8, 773–816.

Haugen, Robert A., and A. James Heins, 1975, Risk and the rate of return on financial assets: some old wine in new bottles, *Journal of Financial and Quantitative Analysis* 10, 775–784.

Hendry, D. F., and H.-M. Krolzig, 2001, *Automatic Econometric Model Selection* (Automatic Econometric Model Selection, London: Timberlake Consultants Press).

Henisz, Witold J, 2000, The institutional environment for economic growth, *Economics & Politics* 12, 1–31.

Henisz, Witold J, 2002, The institutional environment for infrastructure investment, *Industrial and corporate change* 11, 355–389.

Heston, Steven L., and K. Geert Rouwenhorst, 1994, Does industrial structure explain the benefits of international diversification?, *Journal of Financial Economics* 36, 3–27.

Hirshleifer, David, 2001, Investor psychology and asset pricing, *Journal of Finance* 56, 1533–1597.

Hou, Kewei, Chen, and Lu Zhang, 2014, Digesting anomalies: An investment approach, *Review of Financial Studies, forthcoming* .

Jain, Pankaj K, 2005, Financial market design and the equity premium: Electronic versus floor trading, *The Journal of Finance* 60, 2955–2985.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Kang, Jun-Koo, and René M. Stulz, 1997, Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan, *Journal of Financial Economics* 46, 3–28.

Karolyi, G Andrew, 2004, The role of american depositary receipts in the development of emerging equity markets, *Review of Economics and Statistics* 86, 670–690.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lesmond, David A., 2005, Liquidity of emerging markets, *Journal of Financial Economics* 77, 411–452.

Lesmond, David A., Joseph Ogden, and Charles Trzcinka, 1999, A new estimate of transaction costs, *Review of Financial Studies* 12.

Lischewski, Judith, and Svitlana Voronkova, 2012, Size, value and liquidity. Do they really matter on an emerging stock market?, *Emerging Markets Review* 13, 8 – 25.

Liu, Weimin, 2006, A liquidity-augmented capital asset pricing model, *Journal of Financial Economics* 82, 631–671.

Miller, Darius P, 1999, The market reaction to international cross-listings: Evidence from depositary receipts, *Journal of Financial Economics* 51, 103–123.

MSCI, 2011, *MSCI Global Investable Market Indices Methodology*, MSCI Index Research.

Ng, Angela, 2000, Volatility spillover effects from Japan and the US to the Pacific-Basin, *Journal of International Money and Finance* 19, 207–233.

Phylaktis, Kate, and Lichuan Xia, 2006b, Sources of firms' industry and country effects in emerging markets, *Journal of International Money and Finance* 25, 459–475.

Reese, William A, and Michael S Weisbach, 2002, Protection of minority shareholder interests, cross-listings in the united states, and subsequent equity offerings, *Journal of Financial Economics* 66, 65–104.

Reinganum, M. R., 1981, Misspecification of capital asset pricing: Empirical anomalies based on earnings' yield and market values, *Journal of Financial Econometrics* 9, 19–46.

Rouwenhorst, Geert, 1999a, European equity markets and EMU, *Financial Analyst Journal* 57–64.

Rouwenhorst, K. Geert, 1999b, Local return factors and turnover in emerging stock markets, *Journal of Finance* 54, 1439–1464.

Savva, C.S., and C. Aslanidis, 2010b, Stock market integration between new EU member states and the Eurozone, *Empirical Economics* 39, 337–351.

Schotman, Peter C., and Anna Zalewska, 2006, Non-synchronous trading and testing for market integration in Central European emerging markets, *Journal of Empirical Finance* 13, 462–494.

Schwert, G. William, 2003, *Anomalies and market efficiency* (In: Handbook of the Economics of Finance. North Holland, Amsterdam).

Standard & Poor's, 2012, *S&P Dow Jones Indices: S&P Emerging Markets Core Index Methodology.*

Stulz, René M., 1999, Globalization of equity markets and the cost of capital, *Journal of Applied Corporate Finance* 12, 8–25.

Syllignakis, Manolis N., and Georgios P. Kouretas, 2011, Dynamic correlation analysis of financial contagion: Evidence from the Central and Eastern European markets, *International Review of Economics & Finance* 20, 717–732.

van Dijk, Mathijs A., 2011, Is size dead? A review of the size effect in equity returns, *Journal of Banking and Finance* 35, 3263–3274.

Wooldridge, Jeffrey M, 2003, Cluster-sample methods in applied econometrics, *American Economic Review* 133–138.

Zaremba, Adam, and Przemyslaw Konieczka, 2014, Illusionary value, size and momentum premiums in the CEE markets, *Working Paper* .

Tables

Table 3.1: Index Stock Composition and Market Size

This Table reports the composition of both the uncleaned 'Total Sample' and the cleaned 'Liquid Sample' for each of the 16 CEE markets. The sample runs from January 1990 till May 2011. For both the total and liquid sample, we record the number of active (dead) local and preferred stocks, the number of American Depositary Receipts (ADR), Global Depository Receipts (GDR) and direct cross listings (CL), as well as the total number of stocks. Stocks are considered to be liquid if (unadjusted) returns were nonzero in at least 25% of days of the previous year. For the 'Liquid Sample', we report the total number of stocks (Total) as well as the final number of stocks (Final) after replacing illiquid home stocks with the most liquid among the American Depositary Receipts (ADR), Global Depository Receipts (GDR) and direct cross listings (CL) of the same firm. Free-Float reports the average percentage of market capitalization over the sample for which we have free-float data available. The final two columns report the average firm and total market capitalization (in millions of US$) for the year 2010.

	Total Sample				Liquid Sample							
Country	Local	Preferred	ADR/ GDR/CL	Total	Local	Local Preferred	ADR/ GDR/CL	Total	Final	Free-Float (%)	Size 2010 ($mil)	
	Active (Dead)	Active (Dead)	Active (Dead)		Active (Dead)	Active (Dead)	Active (Dead)				Firm Average	Total Market
Bulgaria	374 (39)	-	13 (4)	430	170 (9)	-	-	179	179	61.93	43	3623
Croatia	17 (124)	0 (1)	2 (17)	161	14 (27)	0 (1)	1 (3)	46	41	63.55	690	8964
Czech Rep.	18 (294)	-	22 (79)	413	18 (162)	-	15 (22)	217	182	43.11	6161	43128
Estonia	16 (14)	-	23 (18)	71	15 (13)	-	16 (9)	53	28	46.97	128	1788
Hungary	51 (53)	-	64 (77)	245	39 (34)	-	40 (26)	139	75	39.62	842	28634
Kazakhstan	43 (1)	3 (1)	6 (13)	67	7 (0)	-	0 (2)	9	8	58.33	2662	15972
Latvia	11 (30)	-	32 (1)	74	11 (7)	-	-	18	18	22.71	79	864
Lithuania	36 (33)	-	36 (9)	114	35 (20)	-	0 (2)	57	55	28.35	128	4611
Poland	385 (402)	-	19 (82)	888	350 (124)	-	8 (12)	494	475	39.73	445	149417
Romania	66 (114)	-	0 (2)	182	65 (70)	-	0 (10)	136	135	41.78	291	13683
Russia	524 (256)	25 (87)	144 (146)	1182	243 (76)	2 (11)	98 (69)	499	332	33.12	4343	829480
Serbia	220 (60)	3 (0)		283	32 (0)	1 (0)	-	33	32	95.43	55	1696
Slovak Rep.	76 (32)	-	2 (4)	116	10 (2)	-	2 (2)	16	14	52.33	881	2672
Slovenia	58 (159)	-	1 (0)	218	31 (77)	-	-	108	108	50.48	381	9154
Turkey	340 (95)	-	45 (122)	602	320 (72)	-	17 (40)	449	397	89.47	890	270701
Ukraine	245 (7)	-	15 (43)	310	49 (0)	-	7 (1)	57	52	41.56	915	11796

*Note that for Russia some of the cross-listed stocks are also preferred stocks. **Note that the average size values for the Slovak Rep. are based on the year 2009.

Table 3.2: Summary Statistics

This Table reports summary statistics for the total weekly dollar returns in excess of the 3-month T-Bill rate of our tailor-made indices for each of the 14 CEE countries, of a tailor-made CEE index and of the MSCI CEE, Emerging Market, and Frontier Market indices for the overlapping period January 2002 until May 2011. The table reports the starting year and month for each country/region index, the mean, standard deviation, autocorrelation coefficients as well as the skewness and kurtosis. The starting dates for the tailor-made indices are based on the first date when at least 3 liquid stocks are traded. Stocks are considered to be liquid if (unadjusted) returns were nonzero in at least 25% of days of the previous year. Our tailor-made value-weighted indices include local stocks (from large to small) until 85% of total market capitalization has been reached. The mean and standard deviation are annualized and presented in percent. The column Variance Ratio Test shows the variance of monthly period returns over the variance of 4 overlapping weekly period returns and tests the null hypothesis of the ratio being equal to one which corresponds to a test of standard random walk. Stars (***,**,*) indicate significance at 1%, 5% and 10% for the null hypothesis that the variance ratio is equal to 1.

Country	Starting			Autocorrelation					Variance
	Date	Mean	SD	Lag 1	Lag 2	Lag 3	SK	Kurt	Ratio Test
Bulgaria	1/1/02	44.43	41.25	0.002	0.033	0.131	2.168	24.408	1.527***
Croatia	1/1/97	15.70	25.81	0.016	0.069	0.049	0.916	13.358	1.130
Czech Rep.	2/1/95	28.65	30.29	-0.055	-0.012	0.123	-0.506	6.607	1.073
Estonia	1/1/97	19.94	27.12	0.076	0.140	0.151	-0.027	8.102	1.356***
Hungary	1/1/92	19.36	36.10	-0.018	-0.008	0.102	-0.910	7.854	1.088
Latvia	1/1/99	6.82	25.77	0.004	0.067	0.055	-0.627	7.672	1.232
Lithuania	1/1/99	22.11	25.92	0.111	0.109	0.174	-0.266	8.819	1.436**
Poland	1/1/93	18.42	32.97	-0.067	0.020	0.111	-0.538	6.777	1.112
Romania	1/1/97	31.34	38.65	0.105	0.080	0.212	-0.255	5.437	1.273*
Russia	2/1/95	27.39	40.56	-0.063	-0.071	0.152	-0.516	8.757	1.102
Slovak Rep.	1/1/02	24.09	30.22	-0.044	0.021	0.120	-0.161	7.263	0.920
Slovenia	1/1/97	11.88	23.95	-0.007	-0.001	0.279	-0.458	10.940	1.136
Turkey	1/1/91	22.21	41.33	-0.009	0.012	0.082	-0.357	4.813	1.034
Ukraine*	1/1/06	7.15	35.14	0.050	0.174	0.179	-0.585	6.266	1.403**
Own CEE	1/1/95	22.24	34.48	-0.041	-0.042	0.157	-0.789	7.780	1.078
MSCI CEE	1/1/90	18.31	34.32	-0.068	-0.056	0.166	-0.903	10.080	1.045
MSCI EM	1/1/90	14.29	24.08	0.017	-0.043	0.184	-0.863	9.944	1.128
MSCI FM*	6/1/02	7.54	17.67	0.132	0.207	0.195	-1.791	13.211	1.493***

* Note that the starting dates of Ukraine and MSCI FM for the calculations are later than January 2002.

Table 3.3: Comparison of Tailor-Made Indices with Indices of MSCI, Datastream, and S&P

This Table reports for each country the beta and correlation of the weekly total excess dollar returns on our tailor-made indices with the local and CEE market index returns provided by MSCI, Datastream, and S&P, respectively. The column Starting Date reports the starting year and month. The final three columns also report return correlations among the different index providers. We test the null hypothesis of β *(*ρ*)* being equal to 1, where we use a simple t-test for the regression coefficient and Z-test for the correlation coefficient based on a Fisher transformation $[z = (1/2)\ln((1+\rho)/(1-\rho))]$ for which the asymptotic standard error is $1/\sqrt{T-3}$. Since for the null $\rho = 1$ the Fisher transformation is not valid, we use $\rho = 1 \approx 0.99$ as an approximation. The numbers in brackets record the p-values for different tests.

	MSCI			Datastream			S&P			Correlation		
Country	Starting Date	β	ρ	Starting Date	β	ρ	Starting Date	β	ρ	MSCI/ DS	MSCI/ S&P	DS/ S&P
Bulgaria	5/05	0.67	0.82	10/00	0.74	0.88	12/95	0.60	0.65	0.75	0.69	0.63
	-	(0.00)	(0.00)	-	(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Croatia	5/02	0.63	0.67	-	-	-	4/98	0.34	0.35	-	0.46	-
	-	(0.00)	(0.00)	-	-	-	-	(0.00)	(0.00)	-	(0.00)	-
Czech Rep.	12/94	0.97	0.96	11/93	1.02	0.94	12/94	0.94	0.95	0.95	0.97	0.94
	-	(0.14)	(0.00)	-	(0.21)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Estonia	5/02	0.76	0.87	-	-	-	1/98	0.49	0.59	-	0.63	-
	-	(0.00)	(0.00)	-	-	-	-	(0.00)	(0.00)	-	(0.00)	-
Hungary	12/94	0.94	0.97	6/91	1.02	0.93	12/96	0.95	0.97	0.94	0.99	0.94
	-	(0.00)	(0.00)	-	(0.55)	(0.00)	-	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Latvia	-	-	-	-	-	-	1/98	0.38	0.37	-	-	-
	-	-	-	-	-	-	-	(0.00)	(0.00)	-	-	-
Lithuania	7/08	0.71	0.79	-	-	-	12/95	0.78	0.87	-	0.89	-
	-	(0.00)	(0.00)	-	-	-	-	(0.00)	(0.00)	-	(0.00)	-
Poland	12/92	0.91	0.93	3/94	0.98	0.94	6/95	0.94	0.94	0.98	0.98	0.97
	-	(0.00)	(0.00)	-	(0.26)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 3.3: Comparison of Tailor-Made Indices with Indices of MSCI, Datastream, and S&P (continued)

Country	MSCI			Datastream			S&P			Correlation		
	Starting Date	β	ρ	Starting Date	β	ρ	Starting Date	β	ρ	MSCI/ DS	MSCI/ S&P	DS/ S&P
Romania	11/05	0.91	0.94	12/96	0.99	0.99	1/98	0.64	0.68	0.93	0.70	0.67
	-	(0.20)	(0.00)	-	(0.49)	(0.04)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Russia	12/94	0.91	0.95	1/98	0.61	0.69	12/95	0.78	0.82	0.79	0.90	0.92
	-	(0.00)	(0.00)	-	(0.08)	(0.00)	-	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)
Slovak Rep.	-	-	-	-	-	-	1/97	0.19	0.20	-	-	-
	-	-	-	-	-	-	-	(0.00)	(0.00)	-	-	-
Slovenia	5/02	0.80	0.82	12/98	1.06	0.96	9/02	0.48	0.46	0.85	0.43	0.49
	-	(0.01)	(0.00)	-	(0.08)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Turkey	12/87	0.94	0.85	1/88	0.95	0.86	6/95	0.92	0.85	0.98	0.98	0.99
	-	(0.02)	(0.00)	-	(0.06)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.17)
Ukraine	5/06	0.16	0.21	-	-	-	1/98	0.10	0.14	-	0.74	-
	-	(0.00)	(0.00)	-	-	-	-	(0.00)	(0.00)	-	(0.00)	-
Average	-	0.78	0.81	-	0.92	0.90	-	0.61	0.63	0.90	0.78	0.82
CEE Index	1/88	0.95	0.98	1/94	1.08	0.97	1/06	1.10	0.91	0.98	0.91	0.92
	-	(0.00)	(0.00)	-	(0.00)	(0.00)	-	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)

Table 3.4: Tracking Errors of Tailor-Made Indices, Free-Float Adjusted Indices, and Indices of MSCI, Datastream, and S&P

This Table reports annualized tracking errors (in percent) for our tailor-made indices, free-float adjusted versions of our indices, and indices of MSCI, Datastream, and S&P. For indices a and b, the tracking error is defined as $\sqrt{52}\sqrt{\frac{1}{T}\sum_{t=1}^{T}(r_{at}-r_{bt})^2}$. In the Index vs. Free Float columns, we compare our original tailor-made index (Original) to a tailor-made index that only uses stocks for which the free-float adjusted information is available, but weights them using the full market capitalizations rescaled to make them sum to 100% (FF1) and to a tailor-made index that uses these stocks weighted by their actual float-adjusted market capitalizations (FF2). In the Index vs. Provider columns, we compare our tailor-made index (Index) to indices of MSCI, Datastream and S&P. In the Free Float vs. Provider columns, we compare the free-float adjusted index (FF2) with the commercial indices. In the Provider vs. Provider columns, we compare the provider indices with each other.

Country	Original vs. Free Float			Original vs. Provider			Free Float vs. Provider			Provider vs. Provider		
	Original/FF1	Original/FF2	FF1/FF2	Original/MSCI	Original/DS	Original/S&P	FF2/MSCI	FF2/DS	FF2/S&P	MSCI/DS	MSCI/S&P	DS/S&P
Bulgaria	16.46	16.68	12.57	18.72	15.13	23.45	19.04	17.02	23.46	23.16	24.66	26.68
Croatia	15.31	17.73	9.04	22.33	-	28.57	16.98	-	29.37	-	26.21	-
Czech Rep.	1.95	4.54	4.37	8.02	9.30	7.91	8.72	9.99	8.79	6.13	3.66	6.64
Estonia	0.31	7.52	7.50	15.70	-	27.62	15.69	-	28.49	-	27.38	-
Hungary	0.74	5.26	5.24	8.51	13.07	8.67	7.96	14.80	8.03	14.60	1.38	14.53
Latvia	5.38	11.12	10.50	-	-	21.94	-	-	22.4	-	-	-
Lithuania	2.89	12.58	12.78	23.73	-	18.75	25.83	-	19.47	-	17.60	-
Poland	8.67	9.39	3.41	10.36	9.44	10.11	5.95	4.45	6.50	5.31	6.04	6.04
Romania	7.96	16.42	14.80	14.24	6.35	32.87	12.82	14.96	32.46	15.20	31.88	33.00
Russia	2.65	6.06	5.79	12.29	15.94	11.27	13.97	17.50	12.96	11.57	10.30	13.33
Slovak Rep.	16.37	22.47	19.52	-	-	39.84	-	-	40.65	-	-	-
Slovenia	5.89	6.40	2.34	12.17	5.64	23.41	11.58	5.63	23.91	11.30	24.94	22.59
Turkey	1.34	4.22	4.47	6.28	3.62	5.01	8.36	5.59	6.81	5.33	4.29	4.06
Ukraine	15.51	19.87	17.27	30.17	-	29.42	33.83	-	35.15	-	19.81	-
Average	7.25	11.45	9.26	15.21	9.81	20.63	15.06	11.24	21.32	11.58	16.51	15.86
CEE Index	2.14	5.99	5.59	7.14	9.56	17.19	9.32	11.05	18.79	8.28	17.40	13.90

Table 3.5: Equity Market Development Indicators in 2010

This Table shows the ratio of total market capitalization over GDP; two market liquidity measures, namely (1) turnover, calculated as the ratio of total dollar trading volume per year over the end-of-year market capitalization, and (2) the average percentage of non-zero daily returns, calculated as the yearly average of the value-weighted share of non-zero daily returns (in local currency), and two measures of firm and industry concentration, namely (1) the Herfindahl index and (2) the share of the largest 3 firms (industries). We use FTSE's level-2 classification (10 industries) to calculate the industry measures. All indicators are calculated at the daily frequency, and averaged over the year 2010. The column "Ordinal Rank" ranks the countries from most to least developed, using the aggregated ordinal values of the 5 different measures. The final column reports whether MSCI classifies the market as 'Emerging' or 'Frontier'.

Country	Mcap/ GDP	Liquidity		Concentration				Ordinal Rank	MSCI Classification
		Turn-over	Non-zero returns	Firm HHI	Firm C3	Industry HHI	Industry C3		
Bulgaria	0.075	0.058	0.656	0.029	0.250	0.145	0.707	9	Frontier
Croatia	0.142	0.063	0.900	0.304	0.871	0.293	0.917	11	Frontier
Czech Rep.	0.214	0.494	0.965	0.270	0.900	0.305	0.900	6	Emerging
Estonia	0.093	0.312	0.809	0.101	0.598	0.350	0.882	10	Frontier
Hungary	0.216	1.073	0.954	0.203	0.750	0.161	0.765	3	Emerging
Kazakhstan	0.055	0.007	0.645	0.251	0.881	0.394	1.000	16	Frontier
Latvia	0.035	0.027	0.619	0.140	0.764	0.335	0.890	14	Emerging
Lithuania	0.124	0.069	0.896	0.055	0.380	0.183	0.700	6	Frontier
Poland	0.320	0.372	0.805	0.043	0.309	0.222	0.688	4	Emerging
Romania	0.083	0.138	0.816	0.210	0.684	0.328	0.930	12	Frontier
Russia	0.588	0.813	0.913	0.050	0.331	0.198	0.776	1	Emerging
Serbia	0.049	0.077	0.699	0.071	0.484	0.314	0.898	13	Frontier
Slovak Rep.*	0.029	0.279	0.244	0.586	0.978	0.672	0.799	15	Emerging
Slovenia	0.188	0.087	0.846	0.117	0.523	0.109	0.650	5	Frontier
Turkey	0.386	1.368	0.830	0.029	0.215	0.222	0.729	2	Emerging
Ukraine	0.094	0.097	0.782	0.099	0.694	0.206	0.843	8	Frontier
Germany	0.433	1.14	-	-	-	0.139	0.534	-	Developed
US	1.182	1.76	-	-	-	0.125	0.468	-	Developed

* Note that for the Slovak Republic the values reported pertain to 2009.

Table 3.6: Market Development: Panel Regressions

This Table reports estimation results from multivariate panel regressions of one of the 5 market development indicators (market cap over GDP, turnover, percentage non-zero returns, and Herfindahl measures of firm and industry concentration) on 9 candidate explanatory variables . Our set includes one economic openness measure (ratio of imports plus exports over GDP), three institutional and political reform measures (EBRD transition indicator, Political Constraints measure of Henisz (2000) and a dummy for EU accession announcement), three capital controls and liberalization measures (Financial openness measure of Chinn and Ito (2008), dummies for official equity market liberalization and first sign of liberalization from Bekaert et al. (2005b)) and, lastly, two market specific laws and reforms (dummies for introduction of first prosecution of insider trading law from Bekaert et al. (2005b) and a dummy for introduction of electronic trading system (Jain (2005))). Panel A reports estimation results from panel regressions of each of our market development indicators on all determinants of market development without and with country fixed effects. Panel B reports estimation results for the best multivariate model without and with country fixed effects. Starting with all determinants of market development, we remove all insignificant regressors (if the p-value is above 10%) if their joint F-test is insignificant at a 10% significance level. In case it is significant, we only remove the regressor with lowest absolute t-statistic and run a new model without this regressor. We repeat all previous steps until we are left with only significant regressors. Note that for the turnover ratio neither of the regressors turns out significant. All regressions include country fixed effects and standard errors clustered at country level. Stars (***,**, *) indicate significance at the 1%, 5% or 10% significance level.

Panel A: Multivariate panel regressions

	MCAP/GDP	TO	NONZERO	FIRM HHI	SECTOR HHI	MCAP/GDP	TO	NONZERO	FIRM HHI	SECTOR HHI
Institutional and Political Reforms										
EBRD Transition Indicator	-0.106	-0.128	0.099	-0.087	-0.146	-0.013	2.665	0.249***	-0.333***	-0.528***
	(0.141)	(0.911)	(0.104)	(0.148)	(0.163)	(0.078)	(2.270)	(0.073)	(0.104)	(0.127)
Political Constraint Index	0.042	-7.073*	-0.111	-0.053	-0.214*	0.368***	-7.440*	0.041	-0.068	-0.190
	(0.126)	(3.417)	(0.092)	(0.132)	(0.116)	(0.088)	(3.703)	(0.032)	(0.079)	(0.115)
EU Accession	0.021	-0.366	-0.027	-0.050	-0.015	0.026	-0.694	0.001	-0.005	0.025
	(0.045)	(0.247)	(0.050)	(0.064)	(0.084)	(0.023)	(0.477)	(0.020)	(0.033)	(0.040)
Capital Controls and Liberalization										
Financial Openness	0.006	-0.147	0.017	0.005	-0.012	0.020	-0.448	-0.007	0.009	0.008
	(0.018)	(0.219)	(0.024)	(0.016)	(0.027)	(0.016)	(0.488)	(0.015)	(0.031)	(0.029)
Official Liberalization	0.005	0.069	0.023	-0.043	-0.014	-0.010	0.376	0.041*	-0.008	0.017
	(0.031)	(0.356)	(0.030)	(0.045)	(0.047)	(0.022)	(0.369)	(0.021)	(0.034)	(0.045)
First Sign of Liberalization	0.086	-1.375	-0.016	-0.099	-0.256***	0.047	-1.931	-0.090*	-0.163***	-0.236***
	(0.068)	(1.320)	(0.058)	(0.066)	(0.065)	(0.032)	(1.991)	(0.049)	(0.050)	(0.046)
Market Specific Laws and Reforms										
Insider Trading Law	0.101	0.682	0.158**	-0.018	-0.145*	0.011	-1.206	0.014	0.200*	0.103
	(0.069)	(0.485)	(0.058)	(0.060)	(0.076)	(0.083)	(1.710)	(0.072)	(0.096)	(0.106)
Electronic Trading System	0.075	0.964	0.036	0.031	0.027	0.013	-0.764	0.022	0.026	-0.021
	(0.044)	(0.780)	(0.047)	(0.036)	(0.050)	(0.034)	(0.709)	(0.039)	(0.049)	(0.034)
Economic Openness										
Trade/GDP	-0.047	-0.243	-0.110	0.197***	0.097	0.157*	1.144	0.020	0.077	-0.008
	(0.071)	(0.567)	(0.091)	(0.063)	(0.097)	(0.077)	(1.052)	(0.094)	(0.075)	(0.099)
Constant	0.300	6.413	0.424	0.523	1.278**	-0.293	-0.939	-0.164	1.365***	2.496***
	(0.344)	(4.557)	(0.308)	(0.390)	(0.525)	(0.239)	(3.645)	(0.200)	(0.293)	(0.356)
Country fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
△GDP per Capita	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
△GDP Deflator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.209	0.182	0.268	0.237	0.260	0.686	0.319	0.663	0.592	0.652
Observations	184	184	184	184	184	184	184	184	184	184

Table 3.6: Market Development: Panel Regressions (continued)

Panel B: Best Multivariate Model

	MCAP/GDP	TO	NONZERO	FIRM HHI	SECTOR HHI	MCAP/GDP	TO	NONZERO	FIRM HHI	SECTOR HHI
Institutional and Political Reforms										
EBRD	-	-	-	-0.165**	-	-	-	0.265***	-0.283**	-0.421***
	-	-	-	(0.075)	-	-	-	(0.050)	(0.113)	(0.072)
Political Constraint	-	-6.225*	-	-	-0.240**	0.371***	-5.214**	-	-	-0.194**
	-	(3.290)	-	-	(0.097)	(0.095)	(1.822)	-	-	(0.080)
EU Accession	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-
Capital Controls and Liberalization										
Financial Openness	-	-	-	-	-	0.029***	-	-	-	-
	-	-	-	-	-	(0.009)	-	-	-	-
Official Liberalization	-	-	-	-	-	-	-	0.045**	-	-
	-	-	-	-	-	-	-	(0.015)	-	-
First Sign of Liberalization	0.078***	-	-	-0.102*	-0.264***	0.052**	-	-0.088**	-0.171***	-0.238***
	(0.020)	-	-	(0.053)	(0.049)	(0.017)	-	(0.035)	(0.051)	(0.040)
Market Specific Laws and Reforms										
Insider Trading Law	-	-	0.174***		-0.162***	-	-	-	0.236**	-
	-	-	(0.030)		(0.055)	-	-	-	(0.106)	-
Electronic Trading System	0.089**	-	-	-	-	-	-	-	-	-
	(0.032)	-	-	-	-	-	-	-	-	-
Economic Openness										
Trade/GDP	-	-	-	0.228***	-	0.162**	-	-	-	-
	-	-	-	(0.075)	-	(0.072)	-	-	-	-
Constant	-0.012	5.116*	0.618***	0.658**	0.865***	-0.349***	4.401***	-0.160	1.211***	2.221***
	(0.025)	(2.560)	(0.009)	(0.237)	(0.107)	(0.117)	(1.290)	(0.158)	(0.320)	(0.227)
Country fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
△GDP per Capita	Yes	No	No	No	No	Yes	No	No	No	No
△GDP Deflator	No	No	No	No	No	No	No	No	Yes	Yes
R^2	0.119	0.125	0.162	0.175	0.199	0.679	0.265	0.659	0.579	0.645
Observations	184	184	184	184	184	184	184	184	184	184

Table 3.7: Correlations, Betas and Idiosyncratic Volatilities

This Table reports correlation coefficients (ρ), regression coefficients (β), annualized idiosyncratic volatilities ($\sigma_{i,t}$), their respective trends (trend) and annualized alphas (α) for weekly dollar returns of tailor-made country indices and the CEE index with respect to the benchmark region indices of MSCI, all in excess of the 3-month T-Bill rate. The MSCI benchmark regions are the World, Europe, Emerging Markets, and Russia. Bold numbers stand for significance at the 5% level for correlation and regression coefficients. Idiosyncratic volatilities are calculated based on the residuals from regressions on the benchmarks. We conduct Bunzel and Vogelsang (2005) trend tests using quarterly non-overlapping correlations, betas, and idiosyncratic volatilities (that is, using 13 weekly observations within the quarter) and present the respective annualized trend coefficients. The Bunzel and Vogelsang (2005) test uses a Daniell kernel to maximize the power of this test in small samples. Starts (**, *) indicate significance at the 5% and 10% significance level.

Country	World							Europe						
	ρ	trend	β	trend	$\sigma_{i,t}$	trend	α	ρ	trend	β	trend	$\sigma_{i,t}$	trend	α
Bulgaria	**0.23**	**0.078****	**0.52**	**0.130***	**0.40**	-0.051	**43.05**	**0.26**	**0.074****	**0.47**	0.050	**0.40**	-0.052	**42.39**
Croatia	**0.32**	0.005	**0.53**	-0.004	**0.27**	-0.007	5.64	**0.35**	0.010	**0.48**	-0.004	**0.27**	-0.007	5.08
Czech Rep.	**0.48**	**0.042****	**0.84**	**0.075****	**0.25**	0.000	11.30	**0.54**	**0.041****	**0.79**	**0.054****	**0.24**	-0.001	10.22
Estonia	**0.26**	**0.035****	**0.57**	**0.071****	**0.37**	-0.022	13.34	**0.30**	**0.034****	**0.54**	**0.037***	**0.37**	-0.023	12.69
Hungary	**0.53**	**0.031****	**1.19**	**0.048***	**0.30**	-0.002	10.97	**0.56**	**0.036****	**1.05**	**0.047****	**0.30**	-0.003	9.86
Latvia	**0.21**	**0.035****	**0.39**	0.034	**0.33**	-0.005	14.17	**0.24**	**0.042****	0.37	0.033	**0.32**	-0.004	13.95
Lithuania	**0.30**	**0.034****	**0.43**	**0.043****	**0.24**	0.001	13.65	**0.34**	**0.040****	**0.40**	0.043	**0.24**	0.001	13.41
Poland	**0.42**	**0.037****	**1.09**	**0.034***	**0.38**	-0.020	17.78	**0.45**	**0.040****	**0.97**	0.024	**0.37**	-0.022	16.80
Romania	**0.23**	**0.067***	**0.68**	**0.139****	**0.50**	-0.033	16.22	**0.25**	**0.074****	**0.60**	**0.116****	**0.49**	-0.034	15.57
Russia	**0.37**	**0.046****	**1.36**	**0.106***	**0.57**	**-0.031***	24.81	**0.37**	**0.049****	1.14	**0.075***	**0.57**	**-0.033***	23.71
Slovak Rep.	**0.33**	0.000	**0.53**	0.058	**0.28**	0.019	**26.53**	**0.35**	0.006	**0.46**	0.048	**0.28**	0.018	**25.93**
Slovenia	**0.33**	**0.037****	**0.46**	**0.046****	**0.23**	-0.007	8.60	**0.40**	**0.033****	**0.46**	0.021	**0.23**	-0.007	7.97
Turkey	**0.28**	**0.038****	**1.10**	**0.135****	**0.59**	**-0.027***	7.07	**0.28**	**0.033****	**0.89**	**0.058***	**0.59**	**-0.029***	6.70
Ukraine	0.04	**0.142****	0.08	0.284	**0.35**	-0.010	7.12	0.00	**0.114***	0.01	0.129	**0.35**	-0.009	7.14
Average (all)	0.31	0.045	0.70	0.086	0.36	-0.011	15.73	0.24	0.034	0.52	0.066	0.37	-0.011	15.10
Average (frontier)	0.24	0.057	0.47	0.101	0.34	-0.018	15.37	0.18	0.043	0.33	0.080	0.34	-0.016	14.89
Average(emerging)	0.37	0.033	0.93	0.070	0.39	-0.002	16.09	0.30	0.025	0.71	0.053	0.40	-0.005	15.31
CEE	**0.46**	**0.050****	**1.23**	**0.071***	**0.39**	-0.022	8.43	**0.45**	**0.051****	**1.00**	0.044	**0.40**	-0.024	7.58
Cross-listed	**0.44**	**0.041****	**1.03**	**0.071****	**0.36**	-0.008	15.20	**0.43**	**0.041****	**0.84**	**0.049****	**0.36**	-0.008	14.31
Home-only	**0.40**	**0.035****	**0.55**	0.037	**0.22**	-0.006	6.15	**0.39**	**0.034****	**0.44**	0.020	**0.22**	-0.006	5.68
Large stocks	**0.43**	**0.041****	**0.97**	**0.069****	**0.35**	-0.008	14.19	**0.42**	**0.041****	**0.79**	**0.046****	**0.35**	-0.009	13.35
Small stocks	**0.30**	**0.035****	**0.51**	0.029	**0.27**	-0.005	**22.13**	**0.31**	**0.036****	**0.44**	0.019	**0.27**	-0.006	**21.59**
Liquid stocks	**0.50**	**0.033****	**1.08**	**0.062****	**0.32**	-0.003	**17.28**	**0.49**	**0.034****	**0.89**	0.031	**0.33**	-0.003	**16.31**
Illiquid stocks	**0.20**	**0.026****	**0.35**	0.042	**0.28**	-0.009	**21.93**	**0.20**	**0.024****	**0.28**	0.018	**0.28**	-0.009	**21.64**

Table 3.7: Correlations, betas and idiosyncratic volatilities (continued)

Country	Emerging							Russia						
	ρ	trend	β	trend	$\sigma_{i,t}$	trend	α	ρ	trend	β	trend	$\sigma_{i,t}$	trend	α
Bulgaria	**0.24**	**0.077****	**0.41**	0.120	**0.40**	-0.049	**38.54**	**0.30**	0.045	**0.30**	0.015	**0.39**	-0.048	**37.92**
Croatia	**0.35**	-0.005	**0.42**	-0.032	**0.27**	-0.007	4.16	**0.28**	0.011	**0.15**	0.011	**0.28**	-0.007	3.17
Czech Rep.	**0.56**	**0.035****	**0.71**	**0.039****	**0.24**	0.000	10.38	**0.41**	**0.036****	**0.21**	**0.038****	**0.26**	-0.002	8.25
Estonia	**0.30**	**0.030****	**0.48**	**0.037****	**0.37**	-0.021	11.59	**0.33**	**0.028***	**0.24**	**0.023***	**0.36**	-0.020	9.00
Hungary	**0.55**	**0.036****	**0.89**	**0.047****	**0.30**	-0.003	8.68	**0.46**	**0.042****	**0.30**	**0.054****	**0.32**	-0.004	9.10
Latvia	**0.22**	**0.032***	**0.32**	0.017	**0.33**	-0.004	10.79	**0.14**	**0.045***	**0.10**	0.025	**0.33**	-0.003	11.13
Lithuania	**0.36**	**0.031***	**0.39**	**0.039****	**0.24**	0.002	9.51	**0.26**	**0.038****	**0.14**	**0.032****	**0.25**	0.000	9.31
Poland	**0.49**	**0.037****	**0.92**	0.016	**0.36**	-0.021	15.92	**0.43**	**0.043****	**0.27**	**0.047****	**0.31**	-0.010	3.73
Romania	**0.32**	0.059	**0.68**	**0.090***	**0.48**	-0.033	13.54	**0.24**	0.059	**0.23**	0.044	**0.49**	-0.030	12.21
Russia	**0.47**	**0.045****	**1.25**	0.065	**0.54**	**-0.032***	22.86	-	-	-	-	**-**	-	-
Slovak Rep.	**0.24**	0.011	**0.29**	0.036	**0.29**	0.020	**23.04**	**0.18**	0.011	**0.13**	0.028	**0.29**	0.020	**23.51**
Slovenia	**0.38**	**0.032***	**0.40**	**0.029***	**0.23**	-0.007	7.16	**0.30**	0.024	**0.14**	**0.015***	**0.24**	-0.007	6.36
Turkey	**0.34**	**0.044****	**0.96**	**0.088****	**0.58**	**-0.029***	2.87	**0.34**	**0.040****	**0.34**	0.043	**0.52**	**-0.026***	3.80
Ukraine	0.01	0.092	0.01	0.143	**0.35**	-0.007	7.00	0.04	0.095	0.03	0.105	**0.35**	-0.004	6.77
Average	0.34	0.039	0.58	0.053	0.36	-0.011	13.29	0.29	0.037	0.20	0.033	0.34	-0.009	10.93
Average (frontier)	0.28	0.045	0.40	0.061	0.33	-0.018	13.07	0.25	0.043	0.18	0.035	0.34	-0.017	12.11
Average(emerging)	0.41	0.034	0.76	0.044	0.38	-0.002	13.51	0.33	0.036	0.23	0.039	0.34	0.000	9.76
CEE	**0.56**	**0.053****	**1.09**	**0.050****	**0.37**	-0.025	6.88	**0.54**	0.060	**0.44**	**0.056****	**0.37**	0.28	0.72
Cross-listed	**0.60**	**0.032****	**1.05**	**0.030****	**0.32**	-0.010	11.96	**0.51**	**0.044****	**0.38**	**0.049****	**0.35**	-0.012	6.74
Home-only	**0.53**	**0.027****	**0.54**	0.011	**0.20**	-0.006	4.54	**0.40**	**0.035****	**0.17**	**0.025****	**0.22**	-0.006	2.46
Large stocks	**0.60**	**0.033****	**0.99**	**0.029****	**0.31**	-0.011	11.14	**0.50**	**0.045****	**0.35**	**0.046****	**0.33**	-0.012	6.26
Small stocks	**0.42**	**0.024***	**0.52**	0.000	**0.26**	-0.004	**20.52**	**0.30**	**0.035****	**0.16**	**0.031***	**0.28**	-0.004	**18.78**
Liquid stocks	**0.66**	**0.027****	**1.07**	**0.028****	**0.28**	-0.004	**14.06**	**0.61**	**0.034****	**0.42**	**0.048****	**0.30**	-0.004	7.82
Illiquid stocks	**0.28**	**0.022****	**0.35**	0.008	**0.28**	-0.009	**20.86**	**0.23**	**0.027****	**0.12**	**0.021***	**0.28**	0.66	**19.26**

Table 3.8: Correlation decomposition

This Table shows a decomposition of correlation coefficients for weekly dollar returns of tailor-made country indices and the CEE index with respect to the benchmark region indices of MSCI, all in excess of the 3-month T-Bill rate. The MSCI benchmark regions are the World, Europe, Emerging Markets, and Russia. For each benchmark, we calculate correlations over the last 10 years (ρ_{10}), 2 years (ρ_2), and 1-year ahead (ρ_{+1}) based on the linear trend model from Table 3.7. Note that if the trend coefficient is not significant we use the correlation for the whole sample period (in italics). In the next for columns, we form ratios of the 2-year over the 10-year correlation as well as a model-implied correlation over the 10-year correlation for each benchmark respectively. For the model-implied correlation, we decompose the correlation of a CEE market i with a benchmark b into three main components: the market's beta with respect to the benchmark (β_i^b), the benchmark's return volatility (σ_b^2), as well as country-specific ('idiosyncratic') volatility (σ_i^2) (see also Bekaert et al. (2009c)): $\rho_{t_\beta,t_b,t_i} = \beta_i^b \sigma_b / \sqrt{\left(\left(\beta_i^b\right)^2 \sigma_b^2 + \sigma_i^2\right)}$, where the subscripts indicate the calculation period for beta, the benchmark volatility, and the idiosyncratic volatility.

Country	World							Europe						
	ρ_{10}	ρ_2	ρ_{+1}	$\frac{\rho_2}{\rho_{10}}$	$\frac{\rho_{2,10,10}}{\rho_{10}}$	$\frac{\rho_{10,2,10}}{\rho_{10}}$	$\frac{\rho_{10,10,2}}{\rho_{10}}$	ρ_{10}	ρ_2	ρ_{+1}	$\frac{\rho_2}{\rho_{10}}$	$\frac{\rho_{2,10,10}}{\rho_{10}}$	$\frac{\rho_{10,2,10}}{\rho_{10}}$	$\frac{\rho_{10,10,2}}{\rho_{10}}$
Bulgaria	0.23	0.55	0.69	2.38	1.42	0.96	1.84	0.26	0.58	0.72	2.28	1.31	0.997	1.84
Croatia	0.35	0.20	*0.32*	0.58	0.69	0.97	0.89	0.39	0.25	*0.35*	0.63	0.71	1.004	0.88
Czech Rep.	0.57	0.69	0.74	1.21	0.96	0.97	1.27	0.61	0.74	0.79	1.21	0.93	1.003	1.26
Estonia	0.45	0.48	0.56	1.07	1.21	0.97	0.90	0.49	0.51	0.63	1.04	1.14	1.004	0.90
Hungary	0.64	0.80	0.74	1.24	1.20	0.98	1.06	0.67	0.84	0.83	1.25	1.17	1.003	1.09
Latvia	0.23	0.18	0.43	0.78	0.79	0.96	1.02	0.27	0.24	0.54	0.89	0.86	1.005	1.02
Lithuania	0.35	0.49	0.42	1.40	1.36	0.97	1.07	0.40	0.54	0.53	1.35	1.26	1.004	1.08
Poland	0.62	0.75	0.83	1.21	1.11	0.98	1.13	0.64	0.79	0.89	1.24	1.09	1.003	1.15
Romania	0.42	0.70	0.74	1.68	1.40	0.97	1.30	0.45	0.73	0.82	1.63	1.32	1.004	1.31
Russia	0.62	0.83	0.85	1.33	1.08	0.98	1.29	0.63	0.83	0.87	1.32	1.05	1.003	1.28
Slovak Rep.	0.33	0.43	*0.33*	1.32	1.74	0.97	0.73	0.34	0.44	*0.35*	1.30	1.65	1.005	0.74
Slovenia	0.44	0.57	0.58	1.29	1.07	0.97	1.24	0.48	0.62	0.63	1.27	1.03	1.004	1.24
Turkey	0.54	0.60	0.70	1.11	0.82	0.97	1.31	0.53	0.61	0.68	1.17	0.82	1.004	1.34
Ukraine	-	0.31	0.56	-	-	-	-	-	0.20	0.42	-	-	-	-
CEE	0.69	0.86	0.91	1.86	1.04	0.98	1.22	0.69	0.86	0.93	1.92	1.02	1.003	1.22
Frontier	0.37	0.47	0.53	1.40	1.19	0.97	1.20	0.41	0.49	0.58	1.37	1.13	1.00	1.21
Emerging	0.51	0.61	0.65	1.17	1.10	0.97	1.12	0.53	0.64	0.70	1.19	1.08	1.00	1.13
Cross-listed	0.55	0.60	0.75	1.09	1.01	0.97	1.10	0.55	0.54	0.75	1.34	0.91	1.004	1.07
Home-only	0.52	0.64	0.62	1.22	1.16	0.97	1.08	0.52	0.58	0.65	1.38	1.07	1.004	1.05
Large stocks	0.57	0.62	0.75	1.09	1.03	0.97	1.08	0.56	0.56	0.76	0.99	0.94	1.004	1.05
Small stocks	0.40	0.63	0.61	1.58	1.35	0.97	1.24	0.42	0.60	0.64	1.44	1.22	1.004	1.20
Liquid stocks	0.56	0.62	0.72	1.11	1.04	0.97	1.09	0.56	0.56	0.73	1.29	0.94	1.004	1.05
Illiquid stocks	0.24	0.36	0.44	1.51	1.52	0.96	1.03	0.23	0.32	0.46	1.97	1.35	1.005	1.01

Table 3.8: Correlation decomposition (continued)

Country	Emerging							Russia						
	ρ_{10}	ρ_2	ρ_{+1}	$\frac{\rho_2}{\rho_{10}}$	$\frac{\rho_{2,10,10}}{\rho_{10}}$	$\frac{\rho_{10,2,10}}{\rho_{10}}$	$\frac{\rho_{10,10,2}}{\rho_{10}}$	ρ_{10}	ρ_2	ρ_{+1}	$\frac{\rho_2}{\rho_{10}}$	$\frac{\rho_{2,10,10}}{\rho_{10}}$	$\frac{\rho_{10,2,10}}{\rho_{10}}$	$\frac{\rho_{10,10,2}}{\rho_{10}}$
Bulgaria	0.24	0.48	0.63	2.00	1.38	0.83	1.75	0.30	0.50	*0.51*	1.67	1.15	0.85	1.69
Croatia	0.36	0.14	*0.35*	0.39	0.55	0.84	0.87	0.32	0.21	*0.28*	0.65	0.86	0.86	0.89
Czech Rep.	0.63	0.68	0.73	1.07	0.98	0.88	1.19	0.59	0.63	0.70	1.06	0.96	0.89	1.20
Estonia	0.51	0.42	0.56	0.82	1.10	0.86	0.85	0.44	0.47	0.49	1.07	1.33	0.87	0.90
Hungary	0.66	0.80	0.82	1.21	1.25	0.89	1.05	0.62	0.79	0.93	1.28	1.30	0.89	1.07
Latvia	0.26	0.15	0.42	0.57	0.68	0.84	1.01	0.22	0.10	0.48	0.46	0.53	0.85	1.01
Lithuania	0.43	0.43	0.43	1.01	1.16	0.85	1.00	0.37	0.47	0.47	1.28	1.39	0.86	1.05
Poland	0.69	0.76	0.89	1.11	1.12	0.90	1.08	0.63	0.71	0.92	1.12	1.13	0.90	1.08
Romania	0.50	0.71	*0.32*	1.43	1.34	0.86	1.24	0.45	0.72	*0.24*	1.59	1.44	0.87	1.29
Russia	0.74	0.89	0.95	1.19	1.07	0.91	1.19	-	-	-	1.02	-	-	-
Slovak Rep.	0.23	0.38	*0.24*	1.64	2.45	0.83	0.73	0.16	0.35	*0.18*	2.10	3.15	0.85	0.72
Slovenia	0.51	0.55	0.51	1.08	1.05	0.86	1.17	0.46	0.52	*0.30*	1.13	1.07	0.87	1.19
Turkey	0.64	0.68	0.81	1.06	0.91	0.89	1.24	0.50	0.53	0.87	1.06	0.87	0.88	1.31
Ukraine	-	0.17	0.33	-	-	-	-	-	0.17	0.30	-	-	-	-
CEE	0.81	0.92	1.03	1.13	1.05	0.93	1.14	0.92	0.95	*0.54*	1.03	1.01	0.97	1.04
Frontier	0.42	0.41	0.49	1.12	1.10	0.85	1.15	0.39	0.44	0.46	1.23	1.21	0.86	1.17
Emerging	0.55	0.62	0.70	1.12	1.21	0.88	1.07	0.52	0.58	0.75	1.16	1.28	0.89	1.06
Cross-listed	0.74	0.66	0.83	0.89	0.99	0.91	0.99	0.81	0.66	0.94	0.82	0.95	0.94	0.95
Home-only	0.68	0.66	0.67	0.97	1.08	0.89	0.99	0.58	0.61	0.67	1.05	1.14	0.89	1.03
Large stocks	0.75	0.67	0.83	0.89	1.00	0.91	0.98	0.80	0.66	0.93	0.82	0.96	0.93	0.94
Small stocks	0.48	0.65	0.62	1.36	1.31	0.86	1.20	0.38	0.56	0.61	1.47	1.42	0.86	1.19
Liquid stocks	0.75	0.68	0.80	0.91	1.01	0.91	0.99	0.82	0.68	0.92	0.83	0.97	0.94	0.94
Illiquid stocks	0.29	0.41	0.48	1.40	1.57	0.84	1.03	0.27	0.35	0.46	1.31	1.49	0.85	1.01

Table 3.9: Hurdle Rates

This Table reports correlations between weekly returns on the respective country indices and those on the MSCI World index, the ratios of the return volatilities on the respective country indices and MSCI world, as well as country-specific hurdle rates, again with respect to the MSCI World index. The hurdle rate is the lowest possible expected (excess) return for market i that must be earned for investors with a 100 percent investment in the benchmark to improve their Sharpe Ratio when they invest in market i, given a specific expected (excess) return for the benchmark market. The hurdle rate is defined as $HR_i^b = \rho_{i,b}(E[r_b] - r_f^b)(\sigma_i/\sigma_b) = \beta_{i,b}(E[r_b] - r_f^b)$ with $E[r_b] - r_f^b = 5\%$. We calculate hurdle rates for the whole sample period (HR), i.e. over the the period 2002 to mid 2011, and for the last 2 (HR2) and 5 (HR5) years based on betas from regressions of the returns on the respective country indices on MSCI World index returns (estimated over the full sample, last 2 and 5 years, respectively). The last forward-looking hurdle rate (FHR) is based on a 1-year forward looking beta derived from the linear trend model in Table 3.7. Note that if trend coefficients are not significant, we use a full sample beta (in italics).

Country	World					
	ρ	$\frac{\sigma_i}{\sigma_w}$	HR	HR5	HR2	FHR
Bulgaria	0.23	2.26	2.60	3.88	3.45	4.95
Croatia	0.35	1.42	2.50	2.76	1.77	*2.63*
Czech Rep.	0.58	1.66	4.85	5.38	4.69	6.76
Estonia	0.45	1.49	3.33	4.12	4.10	4.62
Hungary	0.65	1.98	6.42	7.91	8.61	7.37
Latvia	0.29	1.41	2.08	2.71	2.18	1.95
Lithuania	0.35	1.42	2.51	3.37	3.22	2.62
Poland	0.64	1.81	5.75	6.54	6.97	7.16
Romania	0.43	2.12	4.59	6.74	6.99	8.05
Russia	0.64	2.23	7.13	8.53	7.57	9.96
Slovak Rep.	0.33	1.66	2.73	2.85	4.57	*2.65*
Slovenia	0.46	1.32	3.05	3.90	3.17	3.62
Turkey	0.54	2.27	6.15	7.43	5.45	10.48
CEE	0.70	1.89	6.63	7.86	7.07	8.59
US	0.94	0.99	4.69			-
Frontier	0.38	1.67	3.09	4.13	3.78	4.41
Emerging	0.52	1.86	5.02	5.91	5.72	6.62
Cross-listed	0.55	2.09	5.79	7.18	5.88	7.64
Home-only	0.52	1.36	3.55	4.55	4.43	5.56
Large Stocks	0.57	1.87	5.27	6.53	5.52	7.35
Small Stocks	0.40	1.37	2.73	2.79	4.01	3.84
Liquid stocks	0.56	2.05	5.78	7.25	6.17	7.55
Illiquid stocks	0.24	1.48	1.76	1.92	2.77	1.74

Table 3.10: Decomposition of Index Returns into Country and Industry Effects

This Table shows the volatility of the components of the value-weighted weekly country and industry index returns in excess of the CEE market from 1995 until mid 2011. We estimate weekly cross-sectional regressions based on a Heston and Rouwenhorst (1994) model for each return i of country k and industry j: $r_i = \alpha + \beta_1 I_{i1} + \beta_2 I_{i2} + ... + \beta_{10} I_{i10} + \gamma_1 C_{i1} + \gamma_2 C_{i2} + ... + \gamma_{14} C_{i14} + \varepsilon_i$, with industry dummies (I) and country dummies C subject to the constraint that the value-weighted industry and country effects sum to zero. The cross-sectional regressions yield time series of coefficients. The value-weighted country index of country k can be decomposed into a common effect to all countries ($\hat{\alpha}$), a pure country effect ($\hat{\gamma_k}$) , and the value-weighted sum of 10 industry effects ($\sum_{j=1}^{10} \theta_{j,k} \hat{\beta}_j I_{ij}$). The value-weighted industry index of industry j can be decomposed into a common effect to all industries ($\hat{\alpha}$), a pure industry effect ($\hat{\beta_j}$) and the value-weighted sum of 14 country effects ($\sum_{k=1}^{14} \phi_{j,k} \hat{\gamma}_k C_{ik}$). Note that the estimate of α represents the return on the value-weighted CEE index. For each country (industry), we report the annualized volatilities of the value-weighted country (industry) return in excess of the value-weighted CEE index, the pure country (industry) effect and the value-weighted sum of industry (country) effects.

Country	Excess Country Volatility	Volatility of γ_k	Volatility of $\sum_{j=1}^{10} \theta_{j,k} \hat{\beta}_j I_{ij}$
Bulgaria	33.87	25.69	5.13
Croatia	32.03	29.76	6.33
Czech Rep.	27.55	27.89	7.26
Estonia	36.82	33.48	5.87
Hungary	28.17	28.14	5.55
Latvia	32.19	27.26	4.86
Lithuania	29.53	24.52	5.80
Poland	29.60	29.04	5.63
Romania	45.91	42.99	5.67
Russia	35.28	35.56	6.94
Slovak Rep.	36.22	29.04	4.43
Slovenia	43.48	40.21	6.00
Turkey	26.39	26.87	5.60
Ukraine	54.90	30.49	5.03
Average	35.14	30.78	5.72
Industry	Excess Industry Volatility	Volatility of β_j	Volatility of $\sum_{k=1}^{14} \phi_{j,k} \hat{\gamma}_k C_{ik}$
Oil & Gas	19.65	14.97	16.45
Basic Materials	19.39	11.10	15.40
Industrials	18.85	15.17	15.95
Consumer Goods	15.34	7.97	14.01
Health Care	26.07	17.28	20.49
Consumer Service	18.36	12.03	14.11
Telecommunications	26.09	19.83	24.46
Utilities	23.45	16.31	17.27
Financials	16.49	9.32	14.07
Technology	31.24	25.02	18.67
Average	21.49	14.90	17.09

Table 3.11: Cross-Sectional Strategies

This Table reports performance statistics for different characteristics based portfolios. Starting from the first quarter of 1996 until mid 2011, we build portfolios for 2090 liquid stocks in 14 CEE countries based on the following characteristics: 1-quarter lagged size (market capitalization), dividend yield, price earning ratio, and market to book ratio; t-12 to t-1 and t-6 to t-1 months momentum; 1-quarter lagged volatility, low local and global beta (measured over the previous year), and illiquidity (zero returns measure). For each quarter, we rank stocks based on the respective characteristic and construct value-weighted returns for selected stocks in the 25th (L) and 75th quartile (H) as well the H-L /L-H respectively. Based on monthly returns in excess of the 3-month T-bill rate, we calculate means, volatility, and alphas from univariate regressions of the portfolio returns on our own CEE index, MSCI EM index, MSCI Europe index, MSCI World index as well as multivariate regressions on the CEE, Europe (EU), and World (W) indices and on the global Fama-French factors + momentum (Global FF+Mom). We form correlations $\rho_{i,w}$, variance ratios (σ_i/σ_w) and hurdle rates (HR) with respect to the MSCI World index. The last column reports the average liquidity within each portfolio (%NZR). Panel A reports unrestricted portfolios, while Panel B and Panel C report industry and country neutral portfolios. Bold numbers indicate significance at the 5% significance level. Note that for the dividend yield, price-earnings ratio and book-to-market ratio we do not have data for all 2090 stocks but only for 99.7%, 87.1%, and 74% of the original sample, respectively.

Panel A: Unrestricted

Characteristic				(Annualized) alpha with respect to:						Hurdle Rate vis-à-vis World			Liquid
		μ	σ	CEE	EM	Europe	World	CEE-EU-W	Global FF+Mom	$\rho_{i,w}$	$\frac{\sigma_i}{\sigma_w}$	HR	%NZR
Size	L	24.63	32.7	**18.48**	**20.09**	**21.34**	**22.51**	**18.07**	**23.18**	0.47	2.03	4.77	64.3%
(1 quarter lagged)	H	15.91	38.6	**3.24**	9.60	11.26	12.69	**3.17**	11.07	0.61	2.40	7.28	84.2%
	L-H	8.72	33.23	**15.25**	10.50	10.09	9.81	**14.90**	12.12	-0.24	2.06	-	-
dividend yield	L	9.0	32.3	2.86	3.36	4.32	6.01	2.31	-2.67	0.66	2.01	6.67	83.5%
(1 quarter lagged)	H	22.5	37.9	**11.42**	**16.48**	**17.73**	**19.44**	**10.58**	**16.28**	0.60	2.36	7.02	82.2%
	H-L	13.57	30.74	8.56	13.11	13.41	13.42	8.28	**18.95**	0.04	1.91	-	-
P/E	L	29.8	38.7	**19.31**	**23.82**	**25.30**	**26.77**	**19.07**	**18.01**	0.57	2.41	6.90	82.5%
(1 quarter lagged)	H	6.2	40.2	-5.29	0.12	1.74	3.13	-5.31	4.13	0.56	2.50	6.98	79.8%
	L-H	23.61	30.57	**24.60**	**23.70**	**23.55**	**23.64**	**24.39**	13.87	-0.01	1.90	-	-
Market to Book	L	23.2	38.0	14.71	**16.64**	18.99	**20.26**	15.01	12.42	0.55	2.36	6.52	76.5%
(1 quarter lagged)	H	1.8	42.1	**-10.67**	-4.01	-2.82	-1.41	**-10.72**	2.11	0.55	2.62	7.16	83.8%
	L-H	21.40	39.02	**25.38**	20.65	21.81	21.67	**25.74**	10.31	-0.05	2.42	-	-
Momentum	L	20.2	50.0	6.11	12.56	14.99	16.31	7.24	**25.43**	0.56	3.10	8.67	78.4%
(t-12 to t-1)	H	14.7	40.6	2.82	8.84	10.13	11.66	2.32	6.90	0.55	2.52	6.97	83.0%
	H-L	-5.42	37.45	-3.29	-3.72	-4.87	-4.65	-4.92	**-18.53**	-0.15	2.33	-	-
Momentum	L	12.0	51.3	-2.50	4.38	6.34	7.99	-2.14	12.29	0.56	3.19	8.98	78.9%
(t-6 to t-1)	H	16.4	37.6	5.70	10.50	11.95	13.40	5.44	10.53	0.58	2.34	6.78	82.1%
	H-L	4.43	36.15	8.20	6.12	5.61	5.41	7.59	-1.76	-0.20	2.25	-	-
Volatility	L	19.7	28.5	**12.70**	**14.90**	**16.33**	**17.47**	**12.39**	**11.22**	0.56	1.77	4.96	83.2%
(t-3 to t)	H	6.5	48.5	-6.54	-0.98	1.05	2.73	-6.36	7.29	0.57	3.01	8.63	76.2%
	L-H	13.1	37.0	**19.24**	**15.87**	**15.28**	14.74	**18.75**	3.94	-0.32	2.30	-	-

Table 3.11 continued...

Panel A: Unrestricted

Characteristic				(Annualized) alpha with respect to:						Hurdle Rate vis-à-vis World			Liquid
		μ	σ	CEE	EM	Europe	World	CEE-EU-W	Global FF+Mom	$\rho_{i,w}$	$\frac{\sigma_i}{\sigma_w}$	HR	%NZR
low local beta	L	10.0	21.8	6.40	7.11	7.73	8.65	5.69	3.58	0.45	1.35	3.04	70.7%
(1 year lagged)	H	13.9	49.6	-0.69	6.12	8.28	9.87	-0.10	14.99	0.59	3.08	9.09	84.5%
	L-H	-3.9	45.1	7.09	1.00	-0.55	-1.22	5.79	-11.41	-0.43	2.80	-	-
low world beta	L	14.1	29.5	7.61	10.10	10.71	11.98	6.88	9.99	0.52	1.83	4.81	70.6%
(1 year lagged)	H	18.4	42.3	6.79	11.08	13.34	14.84	7.17	13.05	0.61	2.63	8.04	85.7%
	L-H	-4.3	32.9	0.82	-0.98	-2.63	-2.86	-0.29	-3.06	-0.32	2.05	-	-
Illiquidity	L	18.5	35.7	8.37	**12.11**	13.78	**15.28**	8.37	8.90	0.66	2.22	7.29	90.2%
(1 year lagged)	H	22.8	32.4	**15.59**	**19.26**	**20.46**	**21.29**	**15.01**	**14.20**	0.34	2.01	3.44	45.9%
	H-L	4.3	28.9	7.22	7.15	6.68	6.01	6.64	5.30	-0.43	1.80	-	-

Table 3.11 continued...

This Table reports performance statistics for different characteristics based portfolios. Starting from the first quarter of 1996 until mid 2011, we build portfolios for 2090 liquid stocks in 14 CEE countries based on the following characteristics: 1-quarter lagged size (market capitalization), dividend yield, price earning ratio, and market to book ratio; t-12 to t-1 and t-6 to t-1 months momentum; 1-quarter lagged volatility, low local and global beta (measured over the previous year), and illiquidity (zero returns measure). For each quarter, we rank stocks based on the respective characteristic and construct value-weighted returns for selected stocks in the 25th (L) and 75th quartile (H) as well the H-L /L-H respectively. Based on monthly returns in excess of the 3-month T-bill rate, we calculate means, volatility, and alphas from univariate regressions of the portfolio returns on our own CEE index, MSCI EM index, MSCI Europe index, MSCI World index as well as multivariate regressions on the CEE, Europe (EU), and World (W) indices and on the global Fama-French factors + momentum (Global FF+Mom). We form correlations $\rho_{i,w}$, variance ratios (σ_i/σ_w) and hurdle rates (HR) with respect to the MSCI World index. The last column reports the average liquidity within each portfolio (%NZR). Panel A reports unrestricted portfolios, while Panel B and Panel C report industry and country neutral portfolios. Bold numbers indicate significance at the 5% significance level. Note that for the dividend yield, price-earnings ratio and book-to-market ratio we do not have data for all 2090 stocks but only for 99.7%, 87.1%, and 74% of the original sample, respectively.

Panel B: Restricted to have same industry structure

Characteristic				(Annualized) alpha with respect to:						Hurdle Rate vis-à-vis World			Liquid
		μ	σ	CEE	EM	Europe	World	CEE-EU-W	Global FF+Mom	$\rho_{i,w}$	$\frac{\sigma_i}{\sigma_w}$	HR	%NZR
Size	L	26.33	30.4	**19.18**	**21.50**	**22.57**	**23.81**	**19.10**	**21.70**	0.61	1.89	5.72	62.1%
(1 quarter lagged)	H	22.83	38.7	**10.36**	**16.45**	**18.09**	**19.59**	**10.18**	**17.10**	0.61	2.40	7.35	85.0%
	L-H	3.50	26.07	8.82	5.05	4.48	4.22	**8.92**	4.60	-0.20	1.62	-	-
dividend yield	L	17.8	28.5	12.28	**13.11**	**14.03**	**15.39**	**11.87**	6.50	0.62	1.77	5.47	79.9%
(1 quarter lagged)	H	29.8	40.5	**18.15**	**23.52**	**25.16**	**26.73**	**17.64**	**23.70**	0.56	2.51	7.06	80.5%
	H-L	12.02	32.72	5.86	10.41	11.12	11.33	5.77	**17.20**	0.16	2.03	-	-
P/E	L	31.9	33.8	**23.52**	**26.62**	**27.76**	**29.20**	**22.99**	**23.14**	0.58	2.10	6.05	77.9%
(1 quarter lagged)	H	12.9	38.8	1.59	7.22	8.51	9.91	1.35	11.06	0.56	2.41	6.71	79.6%
	L-H	18.99	28.33	**21.93**	**19.39**	**19.24**	**19.28**	**21.64**	12.09	-0.08	1.76	-	-
Market to Book	L	30.8	35.8	**21.93**	**24.89**	**26.83**	**28.03**	**22.22**	**25.76**	0.57	2.22	6.34	79.0%
(1 quarter lagged)	H	9.1	38.3	-1.69	3.47	4.50	6.00	-1.98	6.29	0.58	2.38	6.95	87.2%
	L-H	21.77	28.85	**23.62**	**21.42**	**22.33**	**22.03**	**24.21**	**19.46**	-0.07	1.79	-	-
Momentum	L	28.3	47.1	**15.37**	**21.75**	**23.34**	**24.85**	**15.42**	**28.97**	0.53	2.92	7.76	83.5%
(t-12 to t-1)	H	17.1	36.1	6.73	11.47	12.82	14.19	6.59	10.10	0.58	2.24	6.56	85.4%
	H-L	-11.20	28.68	-8.64	-10.28	-10.52	-10.67	-8.83	**-18.87**	-0.14	1.78	-	-
Momentum	L	13.8	32.3	4.82	8.57	9.63	11.07	4.29	7.03	0.61	2.01	6.09	83.3%
(t-6 to t-1)	H	21.8	44.3	8.75	**15.17**	16.85	**18.32**	8.95	**22.76**	0.57	2.75	7.85	80.7%
	H-L	8.03	27.57	3.92	6.59	7.22	7.25	4.66	**15.74**	0.21	1.71	-	-
Volatility	L	20.3	28.2	**12.70**	**15.57**	**16.78**	**18.02**	**12.19**	10.49	0.59	1.75	5.21	79.8%
(t-3 to t)	H	23.2	46.2	10.85	16.21	18.00	**19.58**	11.11	**23.43**	0.58	2.86	8.24	76.7%
	L-H	-2.9	33.6	1.85	-0.65	-1.22	-1.55	1.08	-12.94	-0.29	2.09	-	-

Table 3.11 continued...

Panel B: Restricted to have same industry structure

Characteristic				(Annualized) alpha with respect to:						Hurdle Rate vis-à-vis World			Liquid
		μ	σ	CEE	EM	Europe	World	CEE-EU-W	Global FF+Mom	$\rho_{i,w}$	$\frac{\sigma_i}{\sigma_w}$	HR	%NZR
low local beta	L	10.9	21.1	7.36	8.04	8.60	9.45	6.91	3.46	0.48	1.31	3.16	69.4%
(1 year lagged)	H	28.2	47.1	**14.20**	**21.04**	**22.65**	**24.44**	**13.93**	**25.57**	0.59	2.93	8.59	84.4%
	L-H	-17.4	42.1	-6.84	-13.00	-14.05	-14.98	-7.02	**-22.11**	-0.42	2.61	-	-
low world beta	L	15.8	26.4	9.87	12.05	**12.69**	**13.79**	9.48	10.64	0.55	1.63	4.52	69.3%
(1 year lagged)	H	20.7	39.4	**9.35**	**13.96**	**15.80**	**17.19**	**9.93**	**17.60**	0.64	2.44	7.84	85.2%
	L-H	-4.9	30.5	0.52	-1.92	-3.11	-3.40	-0.45	-6.96	-0.35	1.89	-	-
Illiquidity	L	18.5	34.3	**8.57**	**12.31**	13.78	**15.31**	**8.42**	10.81	0.67	2.13	7.14	89.9%
(1 year lagged)	H	26.5	26.9	**19.97**	**22.98**	**24.19**	**24.88**	**20.03**	**21.68**	0.44	1.67	3.68	46.5%
	H-L	8.1	23.3	**11.40**	**10.67**	**10.41**	**9.58**	**11.61**	**10.86**	-0.48	1.45	-	-

Table 3.11 continued...

This Table reports performance statistics for different characteristics based portfolios. Starting from the first quarter of 1996 until mid 2011, we build portfolios for 2090 liquid stocks in 14 CEE countries based on the following characteristics: 1-quarter lagged size (market capitalization), dividend yield, price earning ratio, and market to book ratio; t-12 to t-1 and t-6 to t-1 months momentum; 1-quarter lagged volatility, low local and global beta (measured over the previous year), and illiquidity (zero returns measure). For each quarter, we rank stocks based on the respective characteristic and construct value-weighted returns for selected stocks in the 25th (L) and 75th quartile (H) as well the H-L /L-H respectively. Based on monthly returns in excess of the 3-month T-bill rate, we calculate means, volatility, and alphas from univariate regressions of the portfolio returns on our own CEE index, MSCI EM index, MSCI Europe index, MSCI World index as well as multivariate regressions on the CEE, Europe (EU), and World (W) indices and on the global Fama-French factors + momentum (Global FF+Mom). We form correlations $\rho_{i,w}$, variance ratios (σ_i/σ_w) and hurdle rates (HR) with respect to the MSCI World index. The last column reports the average liquidity within each portfolio (%NZR). Panel A reports unrestricted portfolios, while Panel B and Panel C report industry and country neutral portfolios. Bold numbers indicate significance at the 5% significance level. Note that for the dividend yield, price-earnings ratio and book-to-market ratio we do not have data for all 2090 stocks but only for 99.7%, 87.1%, and 74% of the original sample, respectively.

Panel C: Restricted to have same country structure

Characteristic				(Annualized) alpha with respect to:						Hurdle Rate vis-à-vis World			Liquid
		μ	σ	CEE	EM	Europe	World	CEE-EU-W	Global FF+Mom	$\rho_{i,w}$	$\frac{\sigma_i}{\sigma_w}$	HR	%NZR
Size	L	34.29	37.8	**24.88**	**28.99**	**30.22**	**31.50**	**24.88**	**32.32**	0.54	2.35	6.299	59.4%
(1 quarter lagged)	H	23.57	37.9	**11.21**	**17.32**	**18.96**	**20.40**	**11.07**	**18.48**	0.61	2.35	7.169	84.6%
	L-H	10.72	25.90	**13.68**	11.67	11.26	11.10	**13.81**	**13.83**	-0.11	1.61	-	-
dividend yield	L	25.2	39.5	**13.32**	**18.98**	**20.56**	**21.99**	**13.27**	16.86	0.58	2.45	7.162	77.7%
(1 quarter lagged)	H	27.3	39.4	**16.27**	**21.07**	**22.50**	**24.14**	**15.75**	**19.44**	0.59	2.45	7.255	81.7%
	H-L	2.19	22.06	2.95	2.08	1.93	2.15	2.48	2.58	0.01	1.37	-	-
P/E	L	30.4	41.0	**18.58**	**24.17**	**25.58**	**27.21**	**17.98**	**22.26**	0.56	2.54	7.157	80.0%
(1 quarter lagged)	H	15.8	36.9	5.03	10.01	11.66	12.90	5.22	12.35	0.58	2.29	6.635	81.5%
	L-H	14.55	20.71	**13.55**	**14.16**	**13.92**	**14.31**	**12.76**	9.90	0.08	1.29	-	-
Market to Book	L	34.4	41.9	**22.66**	**27.92**	**29.73**	**31.20**	**22.57**	**28.27**	0.56	2.60	7.252	74.5%
(1 quarter lagged)	H	17.1	37.3	6.66	11.55	12.93	14.28	6.41	12.34	0.55	2.32	6.412	85.0%
	L-H	17.30	24.14	**16.00**	**16.36**	**16.80**	**16.92**	**16.16**	**15.93**	0.11	1.50	-	-
Momentum	L	19.8	43.6	7.36	13.21	14.98	**16.37**	7.76	16.66	0.57	2.71	7.733	77.9%
(t-12 to t-1)	H	20.7	38.0	**9.66**	**15.10**	**16.34**	**17.86**	**8.97**	13.43	0.55	2.36	6.546	80.7%
	H-L	0.96	28.23	2.30	1.89	1.36	1.49	1.22	-3.23	-0.14	1.75	-	-
Momentum	L	34.7	40.1	**23.29**	**28.62**	**29.87**	**31.58**	**22.39**	**25.46**	0.56	2.49	7.035	83.4%
(t-6 to t-1)	H	16.4	38.6	4.29	10.40	12.20	13.42	4.52	**14.66**	0.56	2.40	6.765	80.8%
	H-L	-18.26	25.74	**-19.00**	**-18.22**	**-17.67**	**-18.16**	**-17.87**	**-10.80**	-0.03	1.60	-	-
Volatility	L	25.6	36.0	**14.85**	**20.11**	**21.53**	**22.79**	**14.78**	**19.72**	0.57	2.24	6.357	78.9%
(t-3 to t)	H	17.7	41.6	7.03	11.14	12.99	14.59	6.58	12.89	0.55	2.58	7.088	73.6%
	L-H	7.9	28.6	7.82	8.97	8.54	8.20	8.20	6.83	-0.08	1.78	-	-

Table 3.11 continued...

Panel C: Restricted to have same country structure

Characteristic				(Annualized) alpha with respect to:						Hurdle Rate vis-à-vis World			Liquid
		μ	σ	CEE	EM	Europe	World	CEE-EU-W	Global FF+Mom	$\rho_{i,w}$	$\frac{\sigma_i}{\sigma_w}$	HR	%NZR
low local beta	L	21.6	29.7	**14.04**	**17.65**	**18.55**	**19.66**	**13.38**	**18.25**	0.48	1.84	4.441	66.9%
(1 year lagged)	H	22.2	40.5	**9.66**	**15.69**	**17.31**	**18.82**	**9.64**	**19.06**	0.61	2.51	7.627	84.9%
	L-H	-0.6	27.1	4.38	1.97	1.24	0.84	3.75	-0.81	-0.38	1.68	-	-
low world beta	L	18.9	30.0	**10.70**	**14.50**	**15.45**	**16.60**	**10.40**	**15.56**	0.56	1.86	5.222	66.8%
(1 year lagged)	H	20.5	40.4	8.84	**13.91**	15.44	**17.12**	**8.45**	15.28	0.61	2.51	7.652	84.3%
	L-H	-1.6	23.5	1.86	0.59	0.02	-0.51	1.95	0.28	-0.33	1.46	-	-
Illiquidity	L	21.9	38.7	**9.58**	**15.42**	**17.18**	**18.65**	**9.47**	**17.33**	0.61	2.40	7.309	88.2%
(1 year lagged)	H	31.1	37.8	**21.44**	**26.21**	**27.43**	**28.70**	**20.86**	**25.75**	0.47	2.35	5.510	42.3%
	H-L	9.3	25.4	**11.87**	10.79	10.25	10.05	11.39	8.42	-0.23	1.58	-	-

Table 3.12: Comparison with free-floating stocks: Cross-Sectional Strategies

This Table reports performance statistics for different characteristics based portfolios for the original sample and free-float samples in 14 CEE countries for the overlapping period. Starting from the first quarter of 2003, we build portfolios on the full sample of 2,090 stocks (Original) and on a sample of 1,264 stocks for which free-float data is available, either using (rescaled) market weights as before (FF1) or using free-float market weights (FF2). We use the following characteristics: 1-quarter lagged size (market capitalization), dividend yield, price earning ratio, and market to book ratio; t-12 to t-1 and t-6 to t-1 months momentum; 1-quarter lagged volatility, low local and global beta (measured over the previous year), and illiquidity (zero returns measure). For each quarter, we rank stocks based on the respective characteristic from the previous period (from 1-quarter to 1-year) and construct value-weighted returns for selected stocks in the 25th (L) and 75th quartile (H) as well the H-L /L-H respectively. Based on monthly returns in excess of the 3-month T-bill rate, we calculate alphas from multivariate regressions on the CEE, Europe (EU), and World (W) indices and on the global Fama-French factors + momentum (Global FF+Mom). The Original column reports alphas for the original sample, the FF1 column for the free-float stocks with rescaled original market weights, and the FF2 column for the free-float stocks with free-float market weights. Bold numbers indicate significance at the 5% significance level. Note that for the dividend yield, price-earnings ratio and book-to-market ratio we do not have data for all 2090 (1264) stocks but only for 99.7%, 87.1%, and 74% of the original sample, respectively.

Characteristic		(Annualized) alpha with respect to:					
		CEE-EU-World			Global FF+Mom		
		Original	FF1	FF2	Original	FF1	FF2
Size	L	**33.05**	**26.28**	10.39	**33.25**	**24.45**	15.78
(1 quarter lagged)	H	1.59	1.24	1.85	6.50	6.35	7.22
	L-H	**31.46**	**25.04**	8.53	**26.75**	**18.11**	8.56
dividend yield	L	2.33	0.97	-2.17	-0.19	-2.13	-5.87
(1 quarter lagged)	H	**6.54**	**8.03**	**6.89**	10.17	**11.19**	10.00
	H-L	4.20	7.05	9.07	10.37	13.32	15.87
P/E	L	9.85	**10.62**	7.38	14.25	14.63	11.77
(1 quarter lagged)	H	3.08	1.08	4.00	3.88	2.19	5.67
	L-H	6.78	9.53	3.38	10.37	12.44	6.10
Market to Book	L	**14.59**	**14.39**	**14.40**	**19.61**	**19.49**	**20.08**
(1 quarter lagged)	H	-6.44	-7.50	-6.80	-3.09	-4.35	-4.35
	L-H	**21.03**	**21.88**	**21.20**	**22.70**	**23.83**	**24.43**
Momentum	L	15.50	8.55	9.24	**21.40**	**14.64**	**14.59**
(t-12 to t-1)	H	-2.40	-3.89	0.34	1.65	-0.56	3.35
	H-L	-17.90	-12.44	-8.90	**-19.75**	**-15.20**	-11.24
Momentum	L	-0.83	-0.26	-0.41	4.90	5.59	6.45
(t-6 to t-1)	H	3.99	-1.56	-1.32	6.84	1.39	1.46
	H-L	4.82	-1.30	-0.91	1.93	-4.21	-4.99
Volatility	L	**11.43**	**9.98**	**12.10**	**12.43**	**10.06**	12.23
(t-3 to t)	H	-15.01	-8.03	-9.49	-4.60	-0.45	-1.07
	L-H	**26.45**	**18.01**	**21.59**	17.03	10.51	13.30

Table 3.12 continued...

Characteristic		(Annualized) alpha with respect to:					
		CEE-EU-World			Global FF+Mom		
		Original	FF1	FF2	Original	FF1	FF2
low local beta	L	9.35	3.91	4.33	8.14	2.70	3.16
(1 year lagged)	H	0.89	-3.27	-2.72	7.44	3.21	4.21
	L-H	8.46	7.18	7.05	0.70	-0.51	-1.05
low world beta	L	9.21	6.79	6.21	8.52	4.16	3.31
(1 year lagged)	H	2.33	5.41	3.73	8.75	11.30	10.56
	L-H	6.88	1.38	2.48	-0.22	-7.14	-7.25
Illiquidity	L	1.81	1.70	2.07	7.37	7.58	8.35
(1 year lagged)	H	**15.14**	12.73	**14.46**	**18.47**	14.77	**17.23**
	H-L	13.32	11.03	12.39	11.10	7.19	8.88

Table 3.13: Parametric Portfolios with 6 Characteristics

This Table reports estimation results and descriptives (Panel A) as well as performance statistics (Panel B) for (1) our benchmark CEE index and (2) a parametric portfolio policy that is based on six characteristics: size (me) from the previous quarter, market to book ratio (mtbv) from the previous quarter, and momentum for t-12 to t-1 (mom), volatility of daily returns over the previous quarter (vol), percentage of zero (price) returns in local currency over the previous year (illiq), betas from regressions of weekly returns on the value-weighted CEE index (85% threshold) over the previous year (beta). The sample includes monthly returns for 2,090 liquid stocks in 14 CEE countries over the period 1996 until mid 2011. We report results for a relative risk aversion of 5. The first seven columns of Panel A report the estimated coefficients of policies including 3 and 6 characteristics. We also report bootstrapped standard errors in parentheses (1000 replications of draws with replacement from returns sand characteristics series), the percentage of simulated thetas that have the same sign as the estimated one, as well as the simulated thetas' top/bottom deciles. The next column report the average maximum portfolio weight. The last seven columns report the average weighted normalized characteristics of the portfolio $\sum_{i=1}^{N_t} w_{i,t}\hat{x}_{i,t}$. The first two columns in Panel B report the annualized mean and volatility of the monthly portfolio returns in excess of the 3-month T-bill rate. The next columns report alphas from univariate regressions of the portfolio returns on our own CEE index (85% market cap. threshold), MSCI EM index, MSCI Europe index, MSCI World index as well as multivariate regressions on the latter three indices and on the global Fama-French factors + momentum. Bold numbers indicate significance at the 5% significance level. The next two columns report the (annualized) Sharpe Ratio and and (annualized) Information Ratio calculated as the alpha from the regression on our own CEE index (85% threshold) divided by the standard deviations of the residuals from this regression. We form correlations, variance ratios and hurdle rates with respect to the MSCI World index.

Panel A: Estimates and Portfolio Characteristics

	θ_{me}	θ_{mtb}	θ_{mom}	θ_{vol}	θ_{illiq}	θ_{beta}	max w_i (%)	*me*	*mtb*	*mom*	*vol*	*illiq*	*beta*
CEE VW							16.992	1.266	0.156	0.213	-0.091	-0.045	0.496
CEE PW - 3 char.	-3.113	-2.116	1.257				10.241	-0.233	-0.396	0.227			
Std.	(1.818)	(1.709)	(1.396)										
Prob	0.083	0.025	0.120										
10th percentile	-4.825	-5.261	-0.162										
90th percentile	-0.203	-0.963	3.367										
CEE PW - 6 char.	-2.854	-3.684	1.969	1.409	1.503	1.030	8.418	-0.195	-0.416	0.325	0.414	0.327	0.230
Std.	(2.233)	(2.133)	(1.506)	(1.497)	(2.254)	(1.056)							
Prob	(0.128)	(0.005)	(0.053)	(0.204)	(0.119)	(0.110)							
10th percentile	-5.370	-7.627	0.571	-0.705	-0.232	-0.049							
90th percentile	0.291	-2.157	4.302	3.097	5.450	2.626							

Panel B: Performance Statistics

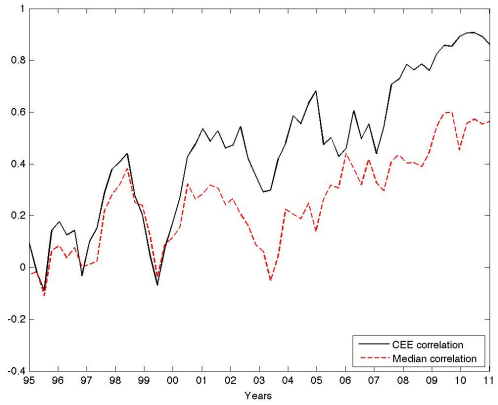
			(Annualized) alpha with respect to:						Sharpe	Information	Hurdle Rate vis-à-vis World		
	Mean (y)	Vol (y)	CEE	EM	Europe	World	CEE-EU-W	Global FF + Mom	Ratio	Ratio (CEE)	ρ	$\frac{\sigma_i}{\sigma_w}$	HR
CEE VW	15.33	37.09	3.75	8.97	10.73	12.13	3.87	11.30	0.41	0.65	0.63	2.30	7.24
CEE PW - 3 char.	23.29	32.50	**13.85**	**17.45**	**18.98**	**20.32**	**13.97**	**19.26**	0.72	1.86	0.66	2.02	6.71
CEE PW - 6 char.	25.31	34.63	**15.27**	**19.25**	**20.82**	**22.19**	**15.47**	**22.28**	0.73	1.91	0.66	2.15	7.05

Figures

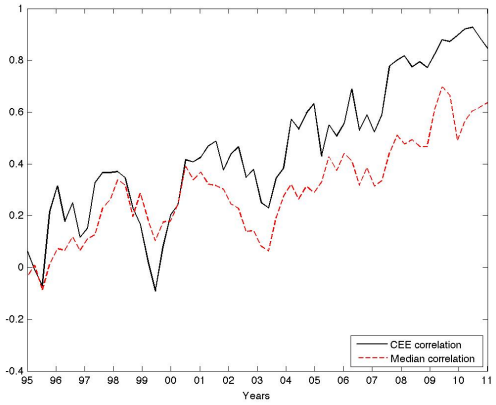
Figure 3.1: The Dynamics of Correlations

The Figures show 4-quarter moving window return correlations of our tailor-made CEE index with returns on 4 benchmarks: MSCI World, MSCI Europe, MSCI Emerging Markets, and MSCI Russia. We also plot at each point in time the (4-quarter moving) median of quarterly return correlations of the individual country indices with each of the 4 benchmarks. All returns are expressed in US dollar and in excess of the 3-month T-Bill rate.

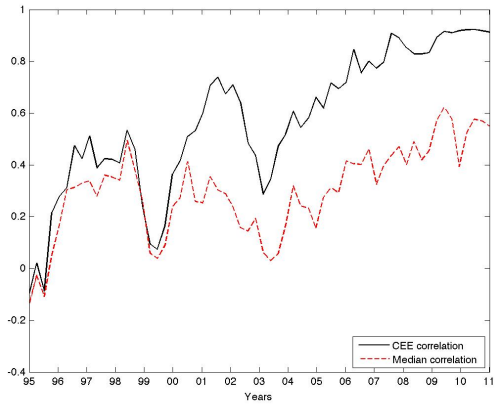
(a) World



(b) Europe



(c) Emerging Markets



(d) Russia

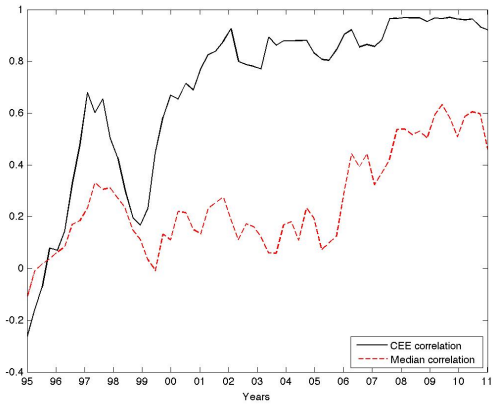
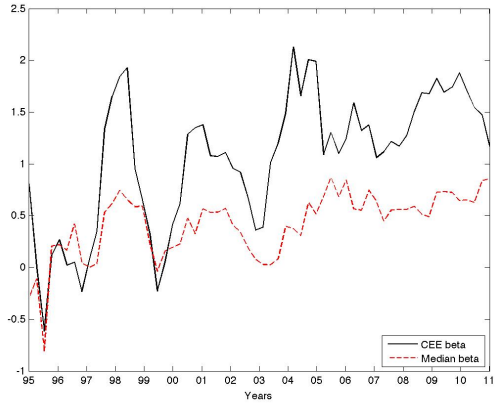


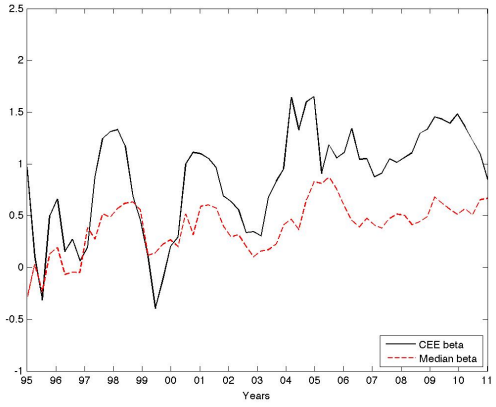
Figure 3.2: The Dynamics of Betas

The Figures show 4-quarter moving averages of quarterly betas (measured over the previous 52 weeks) of our tailor-made CEE index with 4 benchmarks: MSCI World, MSCI Europe, MSCI Emerging Markets, and MSCI Russia. We also plot at each point in time the (4-quarter moving) median of the quarterly betas of the individual country indices with each of the 4 benchmarks. Betas estimates are based on weekly US dollar returns in excess of the 3-month T-Bill rate.

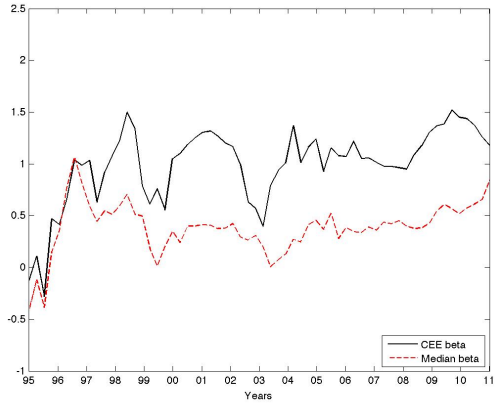
(a) World



(b) Europe



(c) Emerging Markets



(d) Russia

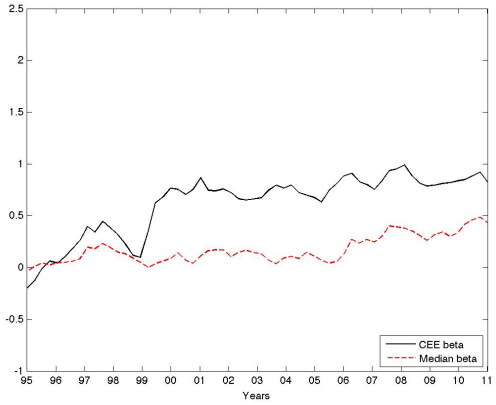
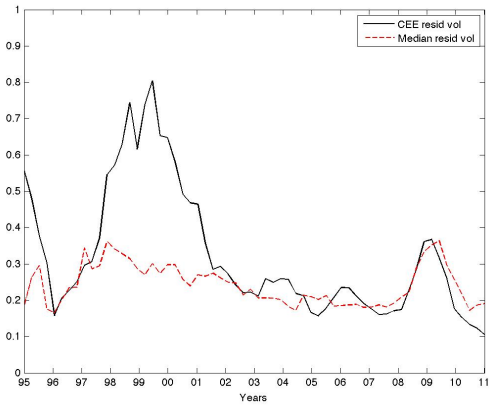


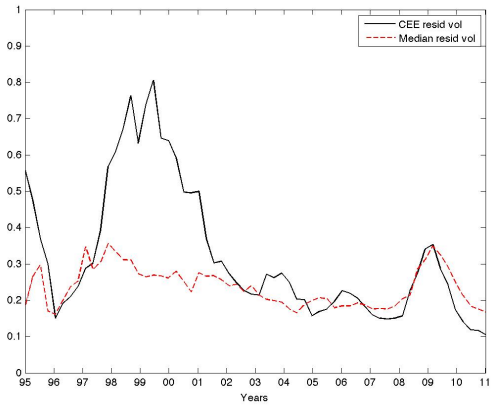
Figure 3.3: The Dynamics of Residual Volatilities

The Figures show 4-quarter moving averages of quarterly (annualized) market-specific returns volatilities for the CEE index as well as the median estimate across all markets. The annualized market-specific volatilities are calculated as $\sqrt{52}$ times the market-specific weekly volatility, which is in turn calculated as the volatility of the residuals from a regression of weekly excess US dollar returns over the past quarter on a specific market index on a constant and the excess US dollar return on one of 4 benchmark returns. The benchmark are: MSCI World, MSCI Europe, MSCI Emerging Markets, and MSCI Russia.

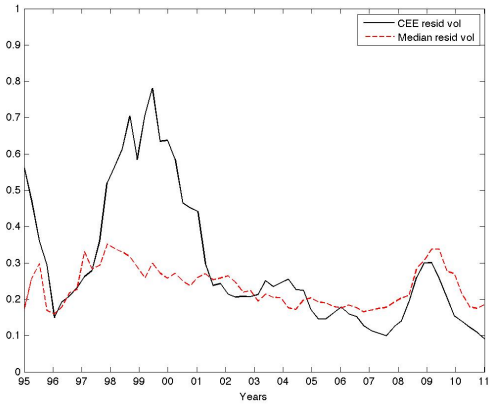
(a) World



(b) Europe



(c) Emerging Markets



(d) Russia

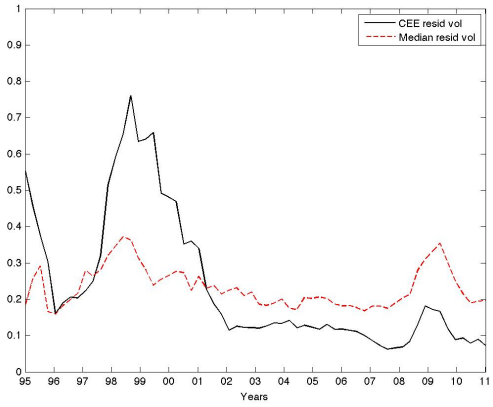


Figure 3.4: Hurdle Rates over Time

The Figures show hurdle rates of weekly dollar returns of CEE country indices and the CEE index with respect to weekly dollar returns of MSCI World index. Hurdle rates are calculated for each quarter over a one year period. For each quarter, the figure shows the median country's hurdle rate as well as the hurdle rate of the CEE index with the world. The hurdle rate is the lowest possible expected (excess) return for market i that must be earned for investors with a 100 percent investment in the benchmark to improve their Sharpe ratio when they invest in market i, given a specific expected (excess) return for the benchmark market. The hurdle rate is defined as $HR_i^b = \rho_{i,b}[E[r_b] - r_f^b](\sigma_i/\sigma_b) = \beta_{i,b}[E[r_b] - r_f^b]$ with $E[r_b] - r_f^b = 5\%$.

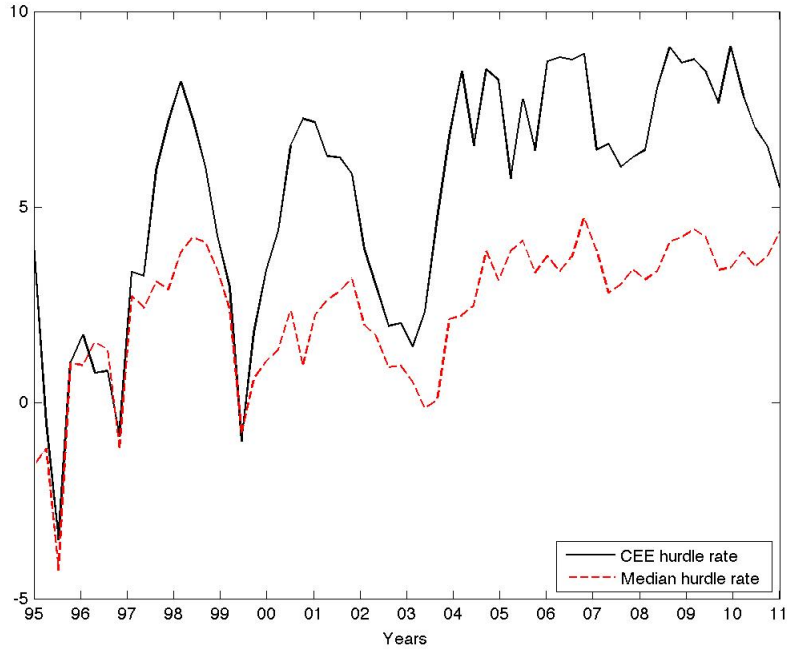


Figure 3.5: Hurdle Rates across Countries

The Figures show hurdle rates of weekly dollar returns of CEE country indices with respect to weekly dollar returns of MSCI World index for the period 2002 to mid 2011. The hurdle rate is the lowest possible expected (excess) return for market i that must be earned for investors with a 100 percent investment in the benchmark to improve their Sharpe ratio when they invest in market i, given a specific expected (excess) return for the benchmark market. The hurdle rate is defined as $HR_i^b = \rho_{i,b}(E[r_b] - r_f^b)(\sigma_i/\sigma_b) = \beta_{i,b}(E[r_b] - r_f^b]$ with $E[r_b] - r_f^b = 5\%$.

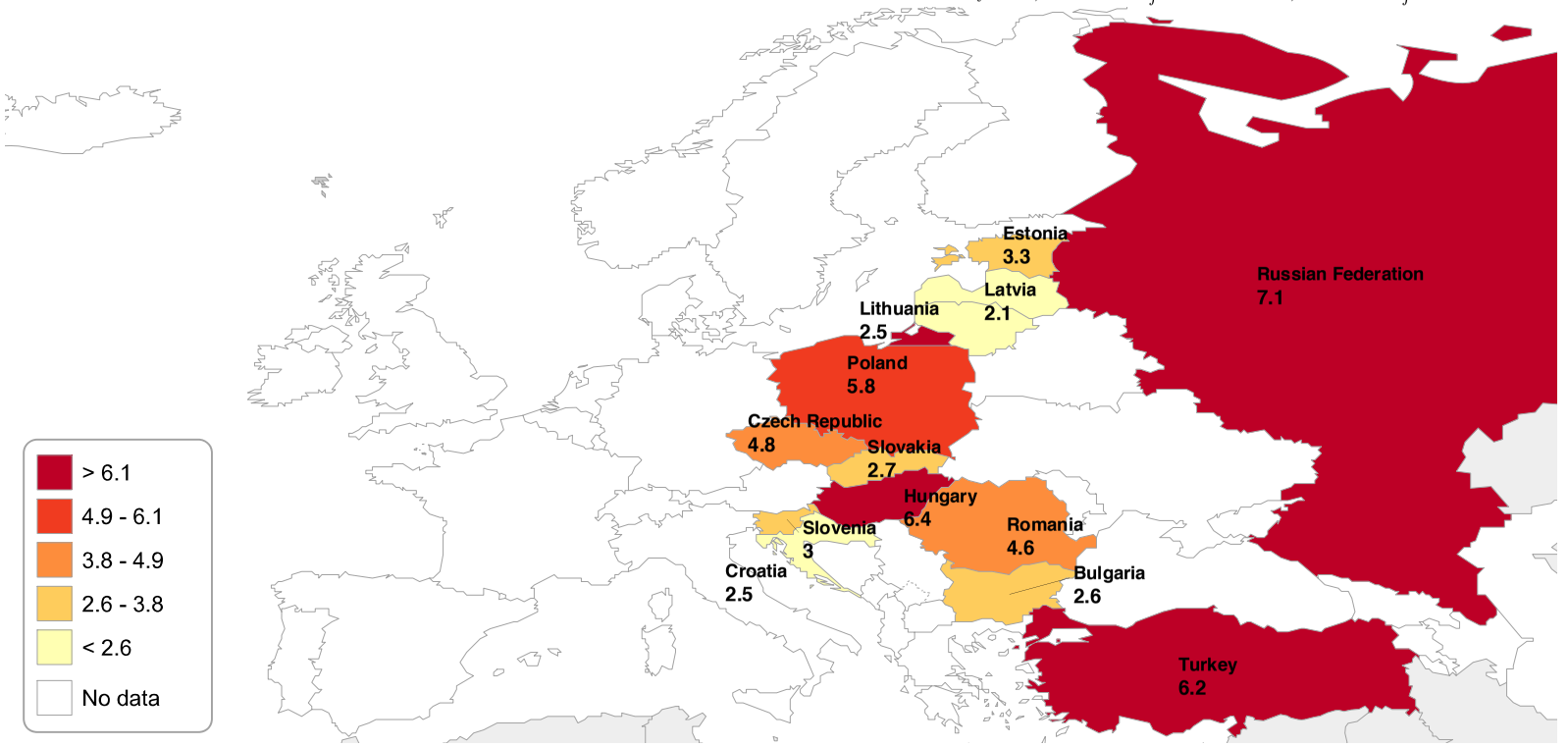


Figure 3.6: Industry and Country Effects

The Figure shows industry and country effects of 2090 weekly stocks returns from all 14 CEE countries. We estimate weekly cross-sectional regressions based on a Heston and Rouwenhorst (1994) model for each return i of country k and industry j: $r_i = \alpha + \beta_1 I_{i1} + \beta_2 I_{i2} + ... + \beta_{10} I_{i10} + \gamma_1 C_{i1} + \gamma_2 C_{i2} + ... + \gamma_{14} C_{i14} + \varepsilon_i$, with industry dummies (I) and country dummies C subject to the constraint that the value-weighted industry and country effects sum up to zero. The cross-sectional regressions yield time series of coefficients. We decompose the R-squared of each regression in a component explained by industry and country effects jointly and individually. In other words, we calculate the variance of the model setting the constant and either the industry or country effect coefficients to zero and divide it by the variance of weekly returns. We take a 1-year moving average of the resulting components.

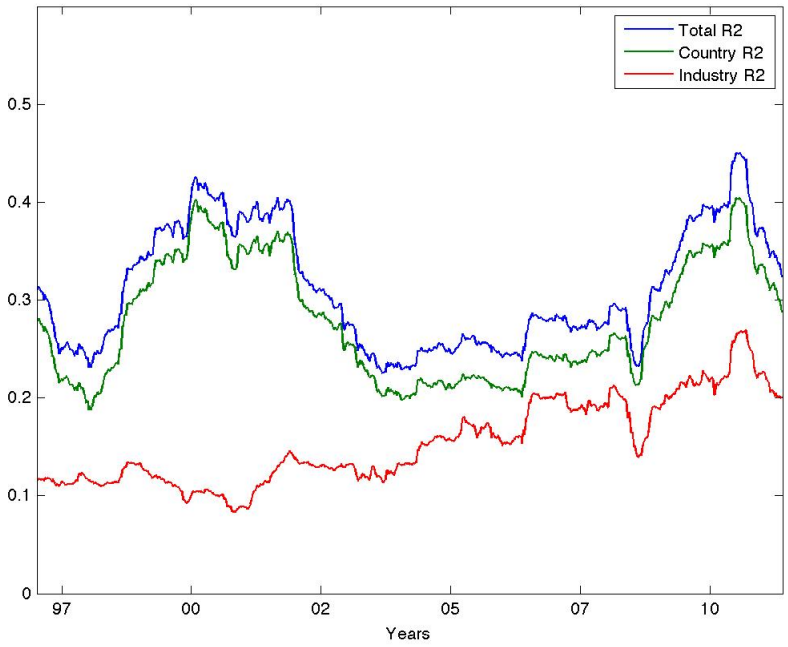
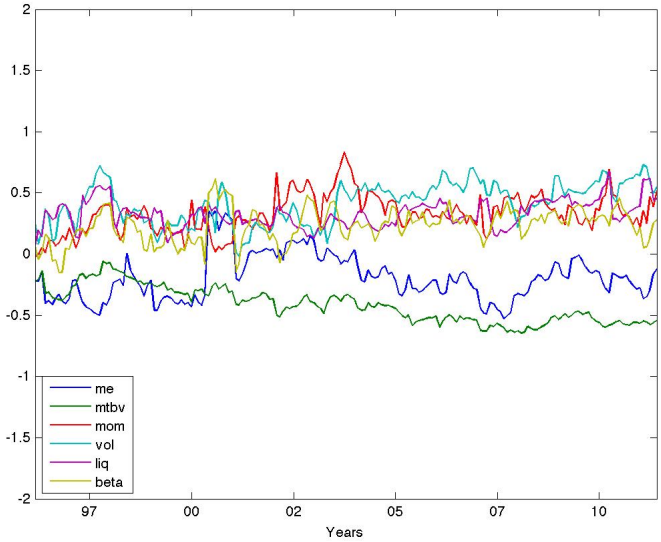


Figure 3.7: Portfolio characteristics over time, $\gamma = 5$

The Figure displays weighted normalized characteristics, $\sum_{i=1}^{N_t} w_{i,t}\hat{x}_{i,t}$, of the optimal portfolio policy with short sale constraints over time for six characteristics: size (*me*) from the previous quarter, market to book ratio (*mtb*) from the previous quarter, and momentum for t-12 to t-1 (*mom*), volatility of daily returns over the previous quarter (*vol*), percentage of zero (price) returns in local currency over the previous year (*illiq*), betas from regressions of weekly returns on the value-weighted CEE index (85% threshold) over the previous year (*beta*). The sample includes monthly returns for 2,090 liquid stocks in 14 CEE countries over the period 1996 until mid 2011. We report results for a relative risk aversion of 5.



A Variable Definitions

Table A.1: Variable Definitions

Variable	Definition
EBRD Transition Indicator	The EBRD Transition Indicator assesses the progress in transition through a set of transition indicators. Progress is measured against the standards of industrialized market economies, while recognizing that there is neither a "pure" market economy nor a unique end-point for transition. The measurement scale for the indicators ranges from 1 to 4+ (coded as 4.3), where 1 represents little or no change from a rigid centrally planned economy and 4+ represents the standards of an industrialized market economy. The reform scores reflect the assessments of EBRD country economists using the criteria described in the methodological notes. The assessment areas are: large scale privatization, small scale privatization, governance and enterprise restructuring, price liberalization, trade and foreign exchange system, competition policy. For the analysis, we calculate an average score of the 6 area scores. Source: EBRD.
Political Constraint Index	The political constraint index is an objective measure of political risk that directly measures the feasibility of policy change based on the a country's political and regulatory structures. The index is based on simple spatial model of political interaction that measures the intensity of constraints a political actor or his replacement faces in his or her choice of future policies. In particular, it identifies the number of independent branches of government (executive, lower and upper legislative chambers, judiciary and sub-federal institutions) with veto power over policy change in each country in every year. This initial measure is then modified to account for the extent of alignment across branches and the heterogeneity or homogeneity of the preferences within each branch. Such alignment increases the feasibility of policy change and thus reduces the level of political constraints. The higher the political constraint index, the more politically constraint a country is and the lower the political risk is (0 to 0.90). A country with no checks and balances would have no constraints on leading politicians, making a veto on key decisions difficult and resembling a totalitarian system (e.g., China has an index of 0). Source: Henisz (2000).
EU Accession	Dummy variable that takes the value of one following the official announcement date of European Union Accession. Source: national websites.

Table A.1 continued...

Variable	Definition
Financial Openness	Index that measures a country's degree of capital account openness (KAOPEN). The measure is based on binary dummy variables that codify the tabulation of restrictions on cross-border financial transactions reported in the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER). In order to measure financial openness the values of the binary variables are reversed, such that the variables are equal to one when the capital account restrictions are non-existent. Then an index of capital "openness" is constructed based on the first standardized principal component of four measures of financial openness. The index takes on higher values the more open the country is to cross-border capital transactions. By construction, the series has a mean of zero. The index has a value of 2.44 for the "most financially open" countries and a value of -1.86 for the "least financial open" countries. Source:Chinn and Ito (2006) and Chinn and Ito (2008).
Official Liberalization	Corresponding to a date of formal regulatory change after which foreign investors officially have the opportunity to invest in domestic equity securities. Official Liberalization dates are based on Bekaert and Harvey (2004), A Chronology of Important Financial, Economic and Political Events in Emerging Markets, `http://people.duke.edu/~charvey/Country_risk/chronology/chronology_index.htm`. This chronology is based on over 50 different source materials. If dates have not been updated or were not available, we added dates for these countries based on information on the official websites of the respective country. For the remaining countries, fully segmented countries are assumed to have an indicator value of zero, and fully liberalized countries are assumed to have an indicator value of one.
First Sign of Liberalization	"First Sign" equity market liberalization dates denote the year associated with the earliest of three dates: Official Liberalization, first American Depositary Receipt (ADR) announcement and first country fund launch. The First Sign indicator takes a value of one on and after the First Sign year, and zero otherwise. First Sign of Liberalization are based on Bekaert and Harvey (2004), A Chronology of Important Financial, Economic and Political Events in Emerging Markets, `http://people.duke.edu/~charvey/Country_risk/chronology/chronology_index.htm`.. This chronology is based on over 50 different source materials. If dates have not been updated or were not available, we added dates for these countries based on information on the official websites of the respective country. As with the Official Liberalization indicator, fully segmented countries are assumed to have an indicator value of zero, and fully liberalized countries are assumed to have an indicator value of one.

Table A.1 continued...

Variable	Definition
Insider Trading Law	Dummy variable that takes the value of one following the introduction of an insider trading law. Source: Bekaert and Harvey (2004), A Chronology of Important Financial, Economic and Political Events in Emerging Markets, `http://people.duke.edu/~charvey/Country_risk/chronology/chronology_index.htm`. This chronology is based on over 50 different source materials. If dates have not been updated or were not available, we added dates for these countries based on information on the official websites of the respective country.
Electronic Trading System	Dummy variable that takes the value of one following the introduction of an insider trading law. Source: Jain (2005) and updates based on national websites.
Trade/GDP	Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product. Imports (Exports) of goods and services represent the value of all goods and other market services received from (provided to) the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments. Source: World Bank.
ΔGDP per Capita	Annual growth of per capita gross domestic product. Source: World Bank.
ΔGDP Deflator	Annual growth of the gross domestic product implicit deflator. Source: World Bank.

B Portfolio Creation

To construct monthly portfolios, we select from all liquid 2,090 stocks of 14 CEE countries starting from the first quarter of 1996 until mid 2011. We start the selection in 1996 in order to be able to cover at least 3 countries and sort on 1-year lagged characteristics. Each quarter, we rebalance the portfolio based on the following characteristics: (i) size based on market capitalization lagged by 1 quarter, (ii) value based on dividend yield (only for 99.7% of all stocks), price earnings ratio (only for 87.1% of all stocks), market to book ratio (only for 74% of all stocks), all lagged by 1 quarter, (iii) momentum based on price return in U.S. dollars for periods from t-12 until t-1 and t-6 until t-1, (iv) volatility of all liquid stocks over the previous quarter (from t-3 until t), (v) betas of regressions of weekly returns on the MSCI World index (global) or the value-weighted CEE index (local) over the previous year, and (vi) illiquidity measured as the percentage of zero (price) returns in local currency for each stock over the previous year.

Based on their rankings, we sort stocks into the top and bottom quartiles and create daily value-weighted indices from which we construct monthly returns. Additionally, we calculate value-weighted daily liquidity indicators measured as the percentage of non-zero returns across stocks in a quartile over time. We rebalance the portfolios on a quarterly basis.

When we restrict portfolios to have the same industry structure, we, first, compute industry weights over time and then split stocks into 10 industries and calculate daily value-weighted industry indices of the top and bottom quartiles for firms in each industry. For each quartile, we combine the industry weights and industry indices into an industry-neutral index. Note that an industry is only considered if it has at least 4 firms which correspond to 4 quartiles, otherwise it is excluded for the respective quarter. Again, we rebalance portfolios on a quarterly basis and calculate the value-weighted percentage of non-zero returns.

When we restrict portfolios to have the same country structure, we, first, compute country weights over time and then split stocks into 14 countries and calculate daily value-weighted country indices of the top and bottom quartiles for firms in each country. For each quartile, we combine the country weights and country indices into a country-neutral index. Note that a country is only considered if it has at least 4 firms which correspond to 4 quartiles, otherwise it is excluded for the respective quarter. Again, we rebalance portfolios on a quarterly basis and calculate the value-weighted percentage of non-zero returns.

C Derivation of Parametric Portfolio Model

As outlined in section 3.5.2, our investor maximizes the following objective function with respect to θ:

$$E\left[r_{p,t+1} - r_{b,t+1}\right] - \frac{\gamma}{2} Var\left[r_{p,t+1} - r_{b,t+1}\right] \tag{A.1}$$

While equation A.1 looks like a mean-variance utility function, the return is the alpha over the benchmark portfolio and the risk is measured relative to the benchmark. In other words, this is a standard tracking error problem. In order to derive the optimal weights we define portfolio weights as follows:

$$w_t = \begin{bmatrix} w_{1t} & w_{2t} & ... & w_{Nt} \end{bmatrix}' ,$$

where w_t is an $(N_t \times 1)$ vector of weights with $i = 1, ..., N$ stocks. We decompose the weights into benchmark weights and an active component that over- or underweights stocks:

$$w_t = w_{bt} + sc_t,$$

where w_{bt} is a $(N_t \times 1)$ vector of stocks' benchmark weights, and sc_t is a $(N_t \times 1)$ vector of scores that is defined as:

$$sc_{i,t} = \frac{1}{N_t} \theta' x_{i,t} \text{ for each stock and } sc_t = (\theta' x_t)' \text{ for all stocks,}$$

where $x_{i,t}$ is a (k $\times$ 1) vector of characteristics of stock i based on observable firm characteristics z ($z = 1, ..., $k), x_t is a (k $\times$ N_t) vector of k characteristics for all N_t stocks, and θ is a time and stock invariant (k $\times$ 1) vector of parameters.

Based on the assumptions above, we define the portfolio and benchmarks returns for the $(N_t \times 1)$vector of returns R_t:

$$r_{p,t+1} = w_t' R_{t+1} = (w_{bt} + sc_t)' R_{t+1},$$

$$r_{b,t+1} = w_{bt}' R_{t+1},$$

We also define the average of the cross-product of characteristics with returns: $\overline{xR} = \frac{1}{T} \sum_{t=0}^{T-1} x_t R_{t+1}$.

We derive the optimal θ by taking the FOC of the sample analog of equation A.1 (setting $N_t = 1$, w.l.o.g):

$$\underset{\theta}{\text{Max}}\ U = E\left[r_{p,t+1} - r_{b,t+1}\right] - \frac{\gamma}{2} Var\left[r_{p,t+1} - r_{b,t+1}\right]$$

Given that $r_{p,t+1} - r_{b,t+1} = \theta' x_{i,t}$ and taking sample analogs, we get:

$$\underset{\theta}{\text{Max}}\ U == \frac{1}{T} \sum_{t=0}^{T-1} \theta' x_t R_{t+1} - \frac{\gamma}{2} \frac{1}{T} \sum_{t=0}^{T-1} \left(\theta' x_t R_{t+1} - \frac{1}{T} \sum_{t=0}^{T-1} \theta' x_t R_{t+1} \right)^2$$

FOC:

$$\frac{\partial U}{\partial \theta'} = \overline{xR} - \frac{\gamma}{T}\sum_{t=0}^{T-1}\theta'(x_t R_{t+1} - \overline{xR})(x_t R_{t+1} - \overline{xR})' = 0$$

$$\overline{xR} - \frac{\gamma}{T}\sum_{t=0}^{T-1}(x_t R_{t+1} - \overline{xR})(x_t R_{t+1} - \overline{xR})'\theta = 0$$

$$\Rightarrow \theta = \frac{1}{\gamma}\left[\frac{1}{T}\sum_{t=0}^{T-1}(x_t R_{t+1} - \overline{xR})(x_t R_{t+1} - \overline{xR})'\right]^{-1}\overline{xR}$$

Q.E.D.